# FORMATIVE FEEDBACK IN A VIRTUAL PATIENT SIMULATOR FOR CLINICAL AUDIOLOGY TRAINING

Lynda May Guard

A thesis submitted in partial fulfilment of the
requirements for the Degree of
Master of Audiology

Department of Communication Disorders
University of Canterbury

March 2013

# Table of Contents

# Acknowledgements

This project was collaborative in nature and in spirit. There are many people I wish to thank. Firstly, I must acknowledge the support of my supervisors, Dr. Catherine Moran and Mr. Jonathan Grady. I thank them for sharing their vision and wisdom, and of course their time. Further thanks are extended to staff at the HITLab NZ, especially Dr. Andreas Duenser and Jakob Zillner. I am especially indebted to Alexandre Heitz, and wish to thank him most sincerely for his support and readiness to assist.

To my MAud colleagues, a resilient and talented bunch of people, I feel truly honoured to have studied alongside you all. Special thanks to Gerard William for his support throughout the thesis process. Further offerings of gratitude go to staff at the Department of Communications Disorders, especially Anne O'Connell for her logistical support in organising participants' incentives (usually at a moment's notice). Also, thank you to the Tertiary Education Committee and GN ReSound for funding assistance, without which this study would not have been possible. Also, I am very grateful that the busy undergraduate CMDS students who were able to find time to participate in this research.

I also wish to thank my work colleagues, friends and my family who have exercised much patience with me, provided food and shelter, editing assistance, emotional support, and motivation to complete what is the peak of my academic pursuits. And finally, I wish to thank my husband Dan, for believing in me and making me his wife in amongst this journey. Onwards together "to unpathed waters, undreamed shores", with all my love and gratitude.

# Abstract

The aim of this study was to examine the effectiveness of different types of feedback on learning pure-tone audiometry using a simulator. Participants were 51 undergraduate students in the Department of Communication Disorders. Participants were randomly allocated to one of two groups whilst undertaking pure-tone audiometry training with the Clinical Audiology Simulator (CAS). One group received summative feedback during the learning task while the second group received formative feedback. Probes were administered to examine participants' knowledge of pure-tone audiometry following training. In addition, a subjective workload analysis was used to measure perceived cognitive load during training and assessment. Between-groups analysis was conducted to establish the effect of feedback on learning and cognitive load. Data regarding how much time each student spent training on the simulator was also collected, and a regression analysis was conducted to evaluate the relationship between time and learning. Formative feedback was found to have a large positive effect on learning in comparison to summative feedback. Cognitive load was perceived as being higher for students receiving formative feedback during training compared to those receiving summative feedback. In subsequent assessment, the formative feedback group reported a lower cognitive load than the summative feedback group. No relationship was observed between time spent training on the simulator and probes outcome. The formative feedback training mode of the CAS proved to be effective in supporting learning and cognitive load in novice students. The findings suggest that the type of feedback employed when using simulators affects learning outcomes for users.

# List of tables

# List of Figures

# List of Abbreviations

| | |
|---|---|
| ASHA | American Speech-Language-Hearing Association |
| BSLP(Hons) | Bachelor of Speech and Language Pathology with Honours |
| CAS | Clinical Audiology Simulator |
| CLT | Cognitive Load Theory |
| CMDS | Communication Disorders |
| CPU | Central Processing Unit |
| HITLab NZ | Human Interface Technology Laboratory New Zealand |
| MAud | Master of Audiology |
| NZAS | New Zealand Audiological Society |
| PhD | Doctor of Philosophy |
| PTA | Pure Tone Audiometry |
| RAM | Random Access Memory |
| SPSS | The Statistical Package for the Social Sciences |
| UoC | University of Canterbury |
| VERG | Virtual Experiences Research Group |
| VP | Virtual Patient |
| VPS | Virtual Patient Simulator |
| VR | Virtual Reality |

# 1 LITERATURE REVIEW

## 1.1 Virtual Patient Simulation

Simulation is widely used in medical education for learning and assessment. It allows medical students and healthcare professionals the opportunity to acquire, refine and rehearse clinical skills (Good, 2003), learn from errors, and increase patient safety (Ziv, Ben-David & Ziv, 2005). Patient simulator technology has its origins in the 1960s and developed rapidly from the 1980s (Rosen, 2008; Bradley, 2006; Good, 2003). Virtual patient-based simulators are recent additions to simulation technology (Botezatu, Hult, Tessma & Fors, 2010), and have been developed to address the increased training expectations in times of reduced training resources (Cook & Triola, 2009). A virtual patient simulator, or VPS, may be defined as an interactive computer programme which offers simulations of real-life clinical scenarios for the purpose of healthcare and medical training, education or assessment (Ellaway & Masters, 2008). A VPS has been developed for use in training and evaluation of clinical skills for Audiology students.

## 1.2 Clinical Audiology Simulator

The development of the Clinical Audiology Simulator (CAS) has been a collaborative project between the Human Interface Technology Laboratory New Zealand (HITLab NZ) and the Department of Communication Disorders, University of Canterbury. Opportunities for beginning Audiology students to gain clinical practice and experience that includes the full scope of Audiology practice in line with theoretic teaching are limited (Duenser, Heitz & Moran, 2010). Foreseeable benefits include students gaining clinical experience before engaging with real clients, including exposure to hearing disorders that occur with low frequency in the normal population, as well as reducing the time required for clinical educators to evaluate student performance (Duenser et al., 2010). The CAS's impact on learning was explored in a pilot study during 2011 as part of Masters and Doctoral research at the University of Canterbury. When utilising the CAS alongside traditional means of clinical training, first year Master of Audiology (MAud) students showed improvements in confidence, perceived level of learning and ability to retrieve information from standardised patients in a clinical role situation (Howland, 2012; Heitz, under review).

The CAS allows Audiology students to practice aspects of the standard diagnostic assessment battery, including case history taking and conducting pure-tone audiometry (PTA) with human-like virtual patients (VPs). Students use the information obtained from the CAS to interpret the responses and plan rehabilitation (Duenser et al., 2010). PTA is a particularly important aspect of hearing evaluation. PTA is used to establish hearing thresholds, a threshold being the quietest level at which a sound is perceived. For conventional PTA, pure-tones are presented across a range of frequencies from 250 to 8,000 Hertz (Hz) (ASHA, 2005). The pure-tones are presented individually at varying intensities (the perceptual equivalent being loudness), measured as decibels (dB). The stimuli are presented to the patient via transducers, typically supra-aural earphones, insert earphones or via bone conduction, selected based on the patient's presentation and requirements (determined from case background information and hearing results as they are found). Further decisions need to be made by the hearing tester such as which ear to start testing first, how loud to start presentations of the stimulus, in what order frequencies are tested, and so on.

A threshold of hearing is established at a given frequency by presenting the pure-tone at varying intensities, seeking the quietest intensity at which it is heard (the patient indicates hearing the sound by clicking a response button). Results are marked on an audiogram, a graph of frequency (x axis) by decibels (y axis). The threshold seeking procedure, as well as other decisions and procedures required when conducting pure-tone audiometry, are informed by clinical protocols, documents guiding the practice of audiology which have been developed by field experts and international organisations to ensure quality and uniformity of practice. Such guidelines include the New Zealand Audiological Society (NZAS) Best Practice Guidelines for Adult Pure-tone Audiometry (2012) and the American Speech-Language-Hearing Association (ASHA) Guidelines for Manual Pure-tone Threshold Audiometry (2005).

## 1.3 Feedback in the Clinical Audiology Simulator

One critical feature of the CAS that requires further examination is the nature of the feedback that is given to users regarding their decisions and performance. Feedback is considered the cornerstone of effective clinical teaching (Cantillon & Sargeant, 2008), and can have significant influence, ideally supporting, optimising and advancing learning (Hattie & Timperley, 2007; Hauer & Kogan, 2012;  Kluger & Denisi 1996; Shute 2008). The definition of feedback involves the transmission of information to an individual following a particular performance, allowing that individual to adjust their performance, or notifying them as to how well they're doing (Chase & Houmanfar, 2009). The transmitter or agent giving feedback may be a teacher, parent, peer, the self, or an experience, where the information transmitted is linked to the performance or understanding of the recipient (Hattie & Timperley, 2007).

The CAS was originally designed with the ability to reveal a model answer once the user completed their assessment with a VP and submitted the worked case. Huxham (2007) defined model answers as ideal, tutor-generated responses to a question that would receive 100% of the marks, and discussed the advantages of providing students with model answers. Model answers are a rapid method of delivering feedback, they avoid personal comments and negative feedback, they require active engagement of students when comparing their responses to the model, and model answers demonstrate the expected standard of work.

In spite of these reported advantages, first year MAud students that were involved in the pilot study felt such feedback was insufficient, and some even failed to identify that they received any feedback at all. Participants were informally asked to provide comments about their experience with the simulator via an anonymous online survey. Their critique of the CAS included comments such as, "there was no feedback so overall it came down to me knowing whether or not I had done it right" and that the "lack of feedback meant you had no idea if the procedure you had used to get to your final answers was right or wrong". They provided suggestions as to what would have been more beneficial, which included increased frequency of feedback provision, such as every time an error was made or at points in the procedure requiring a decision. They also suggested that more helpful feedback would have indicated how to correct an error or provided information about the test protocol, such as via a prompt.

One possible explanation for the students' perceptions that the model answers were not helpful may be that the feedback was summative in nature. Summative feedback is typically delivered at the end of a unit of learning and provides information to the learner about how well they accomplished a given task (White & Weight, 2000). That is, it provides a summary of performance over the entire unit, making it difficult to identify at which point a mistake was made. Summative feedback tends to be outcome and performance orientated which can result in the learner withdrawing from tasks and losing interest in more challenging or difficult tasks (Shute, 2008). Summative feedback also fails to provide the scaffolding a learner requires to achieve a prescribed goal (Chan & Lam, 2010). Furthermore, the learner has the choice as to whether or not they study the worked example. Learners, especially those who may be less experienced, may only study the feedback briefly, engaging less fully in the process of knowledge acquisition (Sweller, Merrienboer & Paas, 1998). When training with the CAS, the goal is to *learn* how to go about the adult diagnostic assessment. Thus, in order to design a feedback intervention for the CAS that more effectively supports learning, it is important to identify what characteristics of feedback lead to effective learning outcomes, and how effective feedback has been implemented in other VPS designs.

## 1.4 Feedback Characteristics

In an attempt to identify the characteristics of feedback which best facilitate the acquisition of clinical skills, literature surrounding feedback in the fields of medicine, education and psychology was reviewed. Very early references to feedback in medical teaching are noted in the writings of ancient Greek physicians, such as Hippocrates (van de Ridder, Stokking, McGaghie, & ten Cate, 2008), and feedback has been studied extensively in the fields of psychology and education with published research dating back to the early 1900s (Chase & Houmanfar, 2009). Whilst it is generally agreed amongst the extensive literature that feedback is an integral part of education, less conclusive evidence exists regarding which characteristics of feedback support learning the most (Shute, 2008). The literature is further confounding in that both feedback and learning are highly complex variables, making it difficult to draw conclusions in order to guide the design of educational systems such as the CAS.

### 1.4.1 Types of feedback

The two feedback types commonly referred to in the literature are verification and elaboration feedback. *Verification feedback* simply lets the recipient know if their response is right or wrong. Verification feedback can be expressed explicitly, by way of a tick or cross, "yes" or "no", or a summative result on a test. Or, it can be expressed implicitly, for example: a student may be allowed to continue uninterrupted in an exercise if going about it in the correct way, or their response yields an expected or unexpected result (Shute, 2008). Anything beyond this is termed *elaboration feedback,* which, in addition to verifying the correctness of a response, also provides information as to why the answer is right or wrong, or offers comments and suggestions with the aim of guiding the recipient in their own revision (Shute, 2008; Chase & Houmanfar, 2009).

Information in the elaborative message can be *specific*, pertaining to each topic, response or action; or *general*, consisting of more conceptual information or definitions (Archer, 2010). Verification feedback is beneficial when compared with no feedback and supports memorisation and retention of less complex learning material, especially in more experienced learners (Kulhavy & Stock, 1989; Shute, 2008). However, verification feedback may only stimulate shallow learning, which is less useful when the learned material is required to be transferred to new contexts, and is less beneficial for novice learners (Kluger & Dinisi, 1996; Shute, 2008). In this respect, the use of elaboration feedback is considered a more constructive approach.

Elaborative feedback (including aspects of both verification and information to guide the learner towards a correct answer) is widely purported to be more effective in supporting learning than verification feedback alone. In their review, Kulhavy & Stock (1989) acknowledged this was "a basic information-processing assumption" (p285), being that error correction would more likely result from more information provided in the feedback message. However, upon reviewing studies conducted up to that time, they discovered a less convincing picture. Only half of the studies showed that more elaborative feedback lead to a higher performance on outcome measures; other studies either showed no significant effects or, somewhat alarmingly, lower post-test scores. They concluded that there was no consistent pattern of results, which reflected the state of research on feedback, i.e. that it was not well understood which feedback characteristics operated most effectively in which situations. They were left wanting for a systematic way to analyse the substance and effect of feedback, an observation that continues to persist 25 years later.

Meta-analyses conducted by Bangert-Drownes, Kulik, Kulik and Morgan (1991) looked at the literature surrounding the effects of feedback that intended to improve retrieval and application of information. They included studies using computer-assisted instruction in their analyses. When considering feedback type, they found that studies showed higher effect sizes for post-test data where learners were exposed to a more elaborative feedback strategy. Conversely, studies showed low average effects where learners were given a more simple verification feedback message. This pattern was observed both within and across studies, and better supported Kulhavy & Stock's (1989) earlier 'assumption', suggesting effective feedback should include elaboration.

Of particular relevance is that the analysis by Bangert-Drownes et al. (1991) looked at studies where outcome measures included the application of information, requiring understanding of the information, acquired with the assistance of feedback. This was a further step from Kulhavy & Stock's (1989) review, which included research using tests requiring retention of the correct answer only, possibly explaining why they found further elaboration had little effect when all that was required was memorisation. Bangert-Drownes et al. (1991) proposed that "elaborate feedback may be more important in the building of conceptual frameworks, drawing of inferences, or applying of rules in complex situations" (p. 234). That is, more elaborative feedback may help the learner to move from surface, fact-based knowledge to deeper conceptual understanding, thereby supporting the recipient more effectively in a test assessing understanding as opposed to simply recalling the correct answer (Butler, Godbole & Marsh, 2013).

### 1.4.2 Feedback timing

Timing of feedback is another variable which has been shown to influence learning outcomes. Timing refers to the point at which the feedback is presented. *Immediate feedback* directly follows a recipient's response or action and is considered to support the development of procedural skills, whereas *delayed feedback* offered after a period of performance promotes skill transfer (Schroth, 1992; Shute, 2008). Kulik and Kulik (1988) conducted a meta-analysis of the literature focussed on the timing of feedback. They found immediate feedback was superior to delayed feedback in tasks that demanded more complex conceptual learning, such as that required in applied situations akin to the clinical environment. *Summative feedback*, also referred to as outcome feedback, typically occurs upon completion of an entire task protocol. The result is a delayed feedback delivery with respect to the difference in time from response made and resulting feedback. While advantages and disadvantages of summative feedback have already been addressed, with specific regard to timing, summative feedback does not lend itself to revision following feedback delivery. This violates a long-standing principle of learning: that the last response made should be the correct one so as to engage learners in active processing (Dempsey & Sales, 1993).

Alternatively, *formative feedback* is "information communicated to the learner that is intended to modify the learner's thinking or behaviour for the purpose of improving learning" (Shute, 2008, p. 1). In contrast to summative feedback, formative feedback is delivered throughout a learning activity. It has been shown to improve students' sense of control over their learning, improving self-efficacy (Chan & Lam, 2007). In addition, it can result in better learning processes whereby the learner persists in the face of adversity, uses more complex strategies and pursues more challenging tasks when acquiring competency in new skills or situations (Shute, 2008).

### 1.4.3 Other feedback variables

Variables in addition to type and timing include: mode of delivery (oral, written, graphic, face-to-face or mediated by other agents, such as via computer-based instruction);

who provides feedback (peer, supervisor, self, expert, machine); who receives it (group, individual); individual differences that can affect the receptiveness to feedback (experience level of the recipient, how the message is interpreted); direction of feedback (to the task (objective) or to the learner (subjective)). Although comprehensive, this is not an exhaustive list of variables. The range of variables alone may confound the effectiveness of feedback and is the likely reason why summary claims regarding the effectiveness of feedback are difficult to make (Bangert-Drownes et al., 1991; Chase & Houmanfar, 2009).

It could be presumed that any feedback is better than no feedback. However, a meta-analysis looking at the effects of verification-based feedback interventions found some results where feedback, compared to no feedback, actually hindered learning and performance (Kluger & Dinisi, 1998). Such disadvantageous variables included feedback that was interpreted as being critical or discouraging, the provision of vague feedback with scores comparing performance of the recipient with peers, and feedback from an external source that delivered feedback orally or interrupted an engaged learner. Also, praise showed negative feedback effects when compared to no feedback.

Although it is difficult to control for individual differences and receptiveness to feedback, disadvantageous types of feedback outlined about feedback should be avoided if the primary aim is to support learning. Hattie & Timperley (2007) summarised characteristics of effective feedback. Specifically, they suggested feedback should be clear, have purpose, be meaningful, and consider students' prior knowledge in order to provide logical connections and prompt active information processing. Shute (2008) provides excellent guidelines regarding the implementation of formative feedback when considering students' prior knowledge, the effects of type and timing of feedback, and negative effects (i.e. what to avoid). These guidelines can be seen in Appendix I.

## 1.5 Feedback in Virtual Patient Simulators

Evidence points towards feedback being a necessary component in Virtual Patient Simulators. In a study that investigated student perceptions of ideal VPS features, feedback was identified among key design principles that facilitated learning and developed clinical reasoning skills (Huwendiek et al., 2009a). Over 100 medical students took part in focus group discussions, where they were interviewed about their experiences with eight different VPs of varying design. Feedback in these systems was based on comparison with expert decision, including elaboration on why a user's decision was right or wrong. The resulting principles of VP design based on participants' responses featured the importance of feedback in a variety of forms, such as receiving specific feedback on all decisions, quantitative formative feedback on overall performance, and explicit guiding of the clinical reasoning process. Students reported that these factors helped them feel "well prepared for dealing with real patients to an extent no other type of teaching had previously achieved" (p. 586). In a separate study, Huwendiek et al. (2009b) consulted with developers to derive a VPS typology in an attempt to bring together the many approaches of VPS design to improve research, development and application of such systems. Their project involved consulting with eight institutions using four different systems spanning six countries to derive a common framework. One of the four categories agreed upon was instructional design, which included feedback as an essential factor. Feedback was very broadly described as "Kinds of feedback, and whether during an activity or at the end (or both)" (p. 745).

Other studies have attempted to identify effective characteristics of feedback in computer simulations. Moreno (2004) reported that feedback of an elaborative nature assisted retention and transfer of knowledge more than feedback that simply indicated right or wrong, especially when novice learners were the recipients of the feedback. In their study, Moreno (2004) aimed to establish the effectiveness of explanatory feedback in helping students learn science from multimedia environments. Students underwent training in a novel subject with a computerised multimedia program and received either elaborative or verification feedback in response to their answers as they worked through an activity. Feedback was delivered orally by the software. The elaborative feedback group produced higher scores on a transfer task, considered the activity more helpful, and found it equally interesting and motivating in comparison to the verification feedback group. These results support the use of elaborative feedback to guide novice students when learning computer-delivered interactive material, as the learner is more effectively facilitated in understanding the material and undergoes deeper learning (Moreno, 2004).

The study by Moreno (2004), in addition to others, contributed to the establishment of design principles that aim to assure educational effectiveness when learning in dynamic, interactive visual environments, including simulations (Moreno & Mayer, 2007; Plass, Homer, & Hayward, 2009). The design principles include feedback as part of the 'principle of guided discovery'. According to this principle, feedback is a method of guiding learning, where novice users in particular learn more effectively in such multimedia contexts. Feedback that both verifies and elaborates on the specific subject being learned, and which consists of information that is easy to interpret, is recommended to best guide novice learners during simulations. These principles suggest that, in order to design an effective simulator (in terms of its educational impact), such feedback should be implemented as a way to support learning, in addition to the consideration of other design principles discussed in these reviews.

Butler et al. (2013) argue that the content of the feedback message alone is the most important aspect of any feedback procedure, irrespective of other variables, such as those discussed by Moreno & Mayer (2007) or Plass et al. (2009). The additional information in elaborative feedback is believed to foster better comprehension of the material, moving the learner from superficial factual knowledge to a more complex understanding of the concept (Butler et al., 2013). In their study conducted with computerised materials, subjects read novel prose passages, were tested on concepts from the texts, and received either verification feedback, elaborative feedback, or no feedback after each question. Subjects returned for a final test two days later consisting of both repeated questions and new inference questions. Those who received verification and elaborative feedback performed equivalently on the repeated questions, and both groups out-performed those who received no feedback. Regarding the new inference questions (which assessed understanding by requiring subjects to transfer their knowledge of the concepts in the texts to a new context), those who received elaborative feedback performed better than those in the verification or no feedback conditions. Whilst providing evidence that more elaborative feedback assists in the application of knowledge, this study did not measure subjects' psychological responses with regard to the learning material, the feedback, or the assessment tasks. Such things as motivation or perceived mental effort are considered to have a major impact on a student's ability to learn (Moreno & Mayer, 2007). By not accounting for such things, they cannot be discounted as having an influence.

Despite the strong argument for elaborative feedback, Mumm & Mutlu (2011) investigated how motivation and persistence were affected by praise, comparative evaluation, and human-like embodiment during computer tasks, which included the comparison of computer-delivered subjective or objective feedback. Subjects that did not receive feedback experienced low motivation, thus the authors concluded “users must have some measure for evaluating their performance in order to sustain intrinsic motivation” (p. 1648). Subjects who received feedback reported increased motivation, regardless of what type. Whilst praise, comparisons and subjective feedback was reportedly unhelpful in traditional settings (Kluger & Denisi, 1996; Shute, 2008), this study was useful in considering such feedback types for implementing in VPS. However, the argument for using a more elaborative type of feedback seems better established.

The above studies and reviews provide support for feedback (mainly elaborative in type) as a VPS design feature, although information that pertains to the effect feedback has on learning in such systems is limited. Drawing conclusions from the literature in order to guide the design of computer-based educational systems, such as the CAS, is difficult considering the complex nature of both feedback and learning. The success of feedback can be affected by many variables such as type, timing, or learner characteristics. Learning may occur at different levels, from surface memorisation and recall, to deeper conceptualisation and application, or may not occur at all if the cognitive demands of the task are too high. To date, more focus has been on the creation of VPS for various immersive learning situations (Cook & Triola, 2009) and more research is required to evaluate the effect that feedback (among other design features) has on learning.

## 1.6 Feedback, Learning and Cognitive Load

It is important to understand the demands placed on a learner during a training task, and what can be done to most effectively support knowledge acquisition. Cognitive demands on learners during problem solving tasks were explored by Sweller (1988), giving rise to Cognitive Load Theory (CLT), an important concept to consider when developing a training tool such as the CAS. Cognitive load (CL) involves the demands placed on working memory (responsible for information processing) during activities that facilitate learning. The three main sources of demand include that which is inherent to the learning task (*intrinsic load*), determined by task or domain complexity. Demand on working memory can also be generated by the design of the learning activity, or the way the information is presented to the learner (*extrinsic load*). Lastly, a challenge to working memory includes the processing requirements for conceptualising, organising, and internalising information, in order for it to be retained and reused (*germane load*).

A basic premise of CLT states that if a learning task is too demanding on working memory, learning will be hindered. It follows that instructional systems be designed to optimise the use of working memory and avoid exceeding the limits of cognitive capacity (Jong, 2009). This may be accomplished by considering the prior knowledge of the learner and avoiding non-essential or confusing information, thus controlling for extraneous intrinsic and extrinsic load, and stimulating processes leading to deep knowledge acquisition, thus increasing germane load. Documented design strategies to avoid cognitive overload during learning tasks are numerous. For example: utilising multiple modalities to deliver information, presenting training material in a simple-to-complex manner, pre-training basic information, learner control of the pacing of information, and the use of feedback (Jong, 2009; Moreno & Meyer 2007; Shute, 2008). Novice or struggling learners in particular can benefit from feedback, as, without guidance, they would easily become overloaded by all the new information (Moreno, 2004; Shute, 2008).

Cognitive load theory draws attention to what happens during learning. As a consequence, it is important to monitor cognitive load when determining the effect feedback has on learning, as different feedback characteristics embedded in the training tool may induce different demands (or relief) on cognitive load. The amount (and quality) of learning that has occurred is commonly measured with transfer tasks, where the learner is asked application questions after a training session (Bransford, Brown & Cocking, 2000). Performance on such tasks can be used as an indirect measurement of cognitive load, as poor results may indicate cognitive load has been too high (Jong, 2009). A more direct method for measuring cognitive load is self-reporting, where perceived effort or workload provides an index of cognitive load (Jong, 2009). Workload is a human-centred rather than a task-centred construct, and involves the “interaction between the requirements of a task, the circumstances under which it is performed, and the skills, behaviours, and perceptions of the operator" (Hart & Staveland, 1988, p. 2). Workload analysis evaluates the cognitive load required to complete the task whilst considering the individual characteristics of the learner. Ideally, effective feedback embedded in the CAS will encourage learning, support the use of working memory and avoid cognitive overload.

## 1.7 Impact of Time on Learning

It is commonly assumed that the more time spent performing a task, the better one will become. It may be proposed that the longer a student spends training in a virtual simulation environment (in this case, with the CAS), the more clinical procedural knowledge is gained. According to principles of deliberate practice, time alone is not the answer. Attaining skill proficiency, including cognitive, psychomotor and clinical skills, involves repetitive performance as well as feedback and assessment (Duvivier et al., 2011; Ericsson, Krampe & Tesch-Romer, 1993). Feedback may be external/behavioural, provided by an expert, guiding the learner to success while reducing errors and frustration (Ericsson et al., 1993), or internal/cognitive, where practice and errors provide feedback to scaffold learning in future attempts (Mayer, 2008). Although time spent on a task may lead to knowledge gains, the addition of feedback is known to promote expertise, where more knowledge may be acquired than a non-feedback condition in the same amount of training time (Bransford et al., 2000). In summary, feedback and time are entwined with learning and must be taken into account when designing an effective training tool.

## 1.8 Research Objectives

In order to address the seemingly deficient characteristics of the existing summative feedback condition in the CAS, an alternative feedback intervention was developed. This involved the ability for the CAS to operate in a training mode, where the program gave formative feedback in response to the user's actions throughout the PTA procedure. The feedback intervention was developed with consideration of the literature and in accordance with Shute's (2008) guidelines of formative feedback (see Appendix I). The study aimed to evaluate the effect of this newly developed formative feedback training mode on learning and cognitive load compared to the original summative feedback condition, by answering the following questions:

1. For students with little prior knowledge of audiology, to what degree will the ability to learn and apply the PTA procedure be influenced by the type of feedback they received during training with the CAS?

2. Will participants' perceived workload during CAS training vary with the type of feedback received, and will this variation persist when applying their knowledge during assessment?

It was hypothesised that training on the CAS with formative feedback will lead to a higher degree of knowledge acquisition and ability to apply this knowledge, reflected by a higher score on a transfer task. Those receiving formative feedback will experience a higher perceived workload when undertaking PTA training with the CAS, but this will reduce during assessment. Those receiving summative feedback will experience the opposite, that is, that perceived workload will be low during training and increase during assessment.

A secondary investigation regarding the effect of time spent training with the CAS on learning outcomes was also conducted. The findings of this study contribute to the development and refinement of the CAS which aims to be integrated into the clinical audiology and associated communication disorders programmes at universities in New Zealand.

# 2 METHODS

## 2.1 Materials

A major component of this study involved the development of the cases and feedback for incorporating into the training mode of the CAS. This section provides information regarding the development history of the CAS and how the resources required for designing the training mode were created.

### 2.1.1 Development of the Clinical Audiology Simulator

The VPS used in this study is referred to as the Clinical Audiology Simulator (CAS), which was developed at the HITLab NZ in conjunction with the Department of Communication Disorders. The CAS is based on a simulation platform initially developed for research study purposes by the University of Florida's Virtual Experiences Research Group (VERG). This group aims to develop experiences with virtual humans for healthcare students and professionals, with a focus on Human–Virtual Patient interactions[1]. This platform was adapted by the HITLab NZ to run as a stand-alone application and then used as a basis for the Immersive Learning Project. Conducted at the HITLab NZ, this project aims to implement a set of Virtual Reality (VR) artefacts for audiologists, medical trainees, and engineering students, with the main intent of evaluating the impact these simulations have on students' learning.

The CAS was designed with the objective to foster communication skills and train students in procedural skills. This takes the form of the standard audiology range of tests including history-taking, pure-tone audiometry, otoscopy, speech audiometry (implementation in progress) and pathology diagnosis. Implementation of these components has involved much collaboration between the software developers and

[1] More information about VERG is available on their website http://verg.cise.ufl.edu/

audiology experts in order to best suit the needs of trainee audiologists and supplement their learning.

### 2.1.2 Development of the CAS Training Mode

The CAS was adapted to include a training mode. Two versions were created with feedback conditions manipulated. For the purposes of this study, participants had their access confined to the pure-tone audiometry (PTA) and otoscopy features of the CAS, as the focus was on supporting procedural learning of PTA, as opposed to taking of case histories or diagnostic abilities. Participants were not required to identify the likely cause of the audiological complaint. Their objective was to undertake the procedure of conducting a hearing test with each virtual patient.

#### 2.1.2.1 Case development

Four cases were developed for use as Virtual Patients (VP) in the training mode. Case demographics, background information, a relevant otoscopy image, a model answer audiogram and notes regarding the hearing testing procedure were provided to the developer for the creation of each VP (see Appendix II). These cases involved typical audiological aetiology and presentations:

- Case One had outer ear infection (otitis externa) and normal hearing,
- Case Two had a mild, noise-induced hearing loss,
- Case Three had a low frequency conductive hearing loss due to Otosclerosis,
- Case Four had a unilateral, idiopathic, sudden sensorineural hearing loss.

The cases, when worked through in order from one to four, were progressively more challenging in terms of PTA (including masking) requirements, judged on the basis of the order in which these concepts are taught in traditional clinical audiology training and presented in protocol documents. Furthermore, each case was designed to target particular concepts of the Clinical Audiology Protocols:

- Case One focussed on the basics of hearing testing, such as transducer selection and the modified Hughson-Westlake threshold seeking procedure (Carhart & Jerger, 1959),
- Case Two built upon these aspects and introduced inter-octave frequency testing,
- Cases Three and Four required additional bone conduction testing and masking on different levels of complexity.

This graduated introduction of hearing testing concepts aimed to support the acquisition of foundation audiometry skills and knowledge, to not overwhelm or discourage novice learners early on, whilst also covering more difficult concepts to ensure that the simulator was a valid tool for clinical audiology training. The four cases developed for use in the training mode were checked and approved by an expert Clinical Educator in audiology (the Clinical Co-ordinator of Audiology at the Department of Communication Disorders) before being programmed into the software.

#### 2.1.2.2 Feedback development

Two versions of the training mode were created with the same four cases, one version with summative feedback, and one with formative feedback. The type of feedback participants received during training with the CAS served as the independent variable of this study.

**Version One: Summative feedback condition**

The CAS was originally developed to give feedback in the form of revealing the VP's actual audiogram once the user submitted their audiogram obtained. The summative feedback training mode of the CAS adapted for this study preserved this design, revealing the model answer audiogram alongside the participant’s findings for comparison (see Figure 1). This screen appeared once the participant deemed they had completed the hearing test and selected the ‘Submit Results’ tab. The participant was able to visually compare the audiogram they obtained against the model answer, but no other information was offered about accuracy (regarding procedure or results), nor were participants able to go back and make any changes to their submitted results.

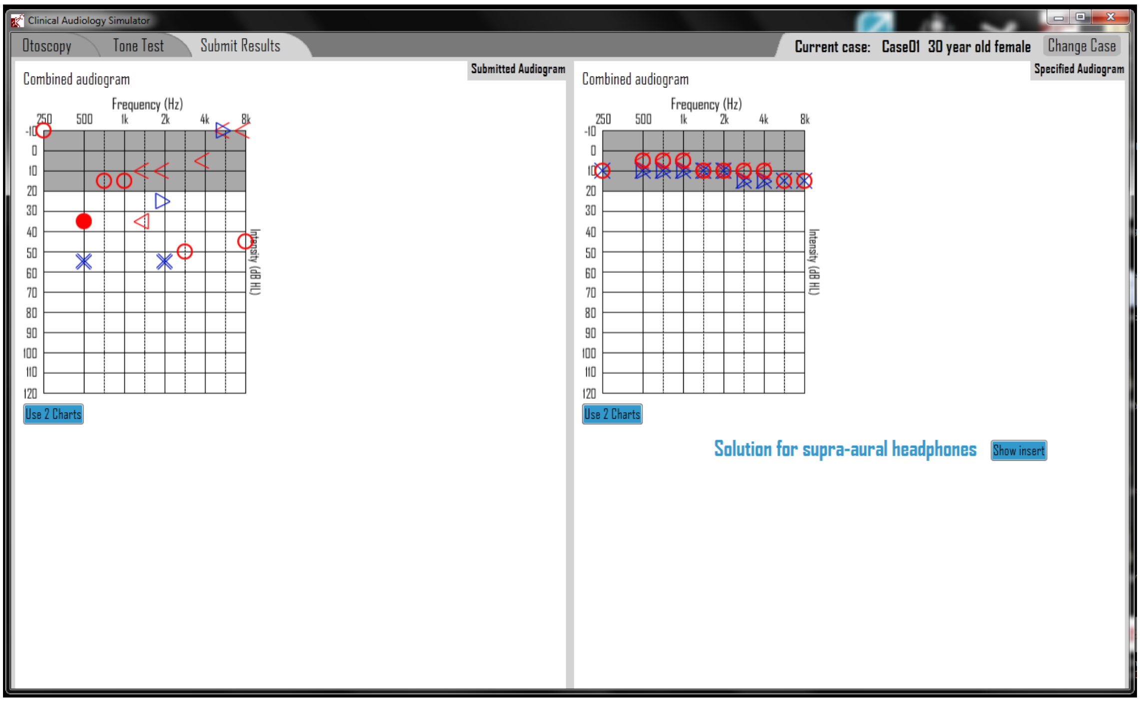



**Figure 1.** *Screen shot of model answer provided after submitting audiogram obtained (on the left) showing model answer (on the right) for Case One*

### Version Two: Formative feedback condition

A second version of the CAS was designed with formative feedback, which took the form of on-screen messages that would appear in pop-up windows during the hearing testing procedure in response to the user's actions. The feedback messages were provided to the software developers as a set of rules to guide the correct execution of the PTA procedure (see Appendix III). This set of rules was based primarily on the University of Canterbury (UoC) Speech and Hearing Clinic Audiology Protocols and Guidelines (unpublished; 2012), and supplemented where necessary by the New Zealand Audiological Society (NZAS) Best Practice Guidelines for Adult Pure-tone Audiometry (2012), and the American Speech-Language-Hearing Association (ASHA) Guidelines for Manual Pure-tone Threshold Audiometry (2005). Again, the rules were checked and approved by the Clinical Co-ordinator of Audiology prior to programming.

Different types of feedback messages were implemented in the design of the formative feedback version of the CAS training mode. These messages included elaborative feedback, which employed implicit verification as messages would pop up only when the user made an error (Figure 2). The elaborative component provided information regarding the protocol at that particular step of the PTA procedure, from which the user could extract the correct response or action (based on the feedback rules seen in Appendix III). A second type of feedback message included positive reinforcement, serving to let the user know they had completed an aspect of the testing procedure to sustain task motivation (Figure 3). A further type of feedback message consisted of hints triggered by a user's action that would be considered inefficient (Figure 4). In addition to these feedback messages, a model answer would also appear once the audiogram obtained was submitted (Figure 5).

### 2.1.3 Equipment

Both training versions of the CAS were deployed on a total of five computers (four PCs and one laptop) located at the HITLab NZ. Minimum system requirements for running the CAS include a Windows XP Operating System, a 2.3 GHz Central Processing Unit (CPU), 2 GB of Random Access Memory (RAM), and a PCI-E Graphics Card with 256MB RAM.

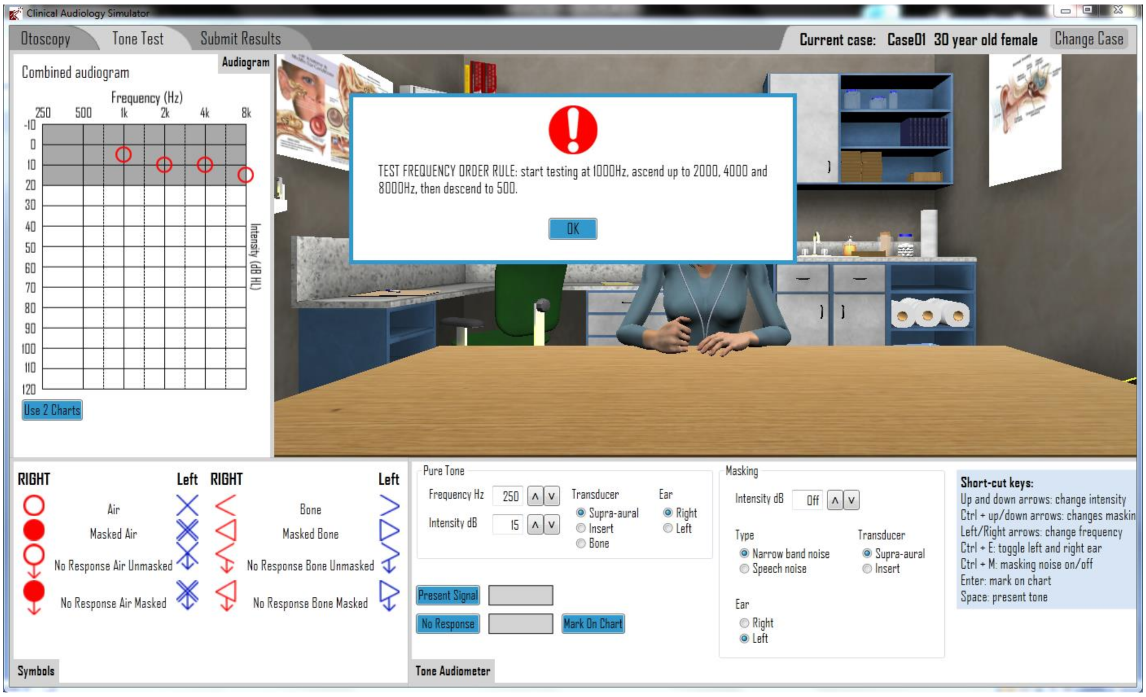



**Figure 2.** *Screen shot of elaborative feedback message in response to an incorrect attempt to test hearing outside of the stipulated testing order of frequencies*

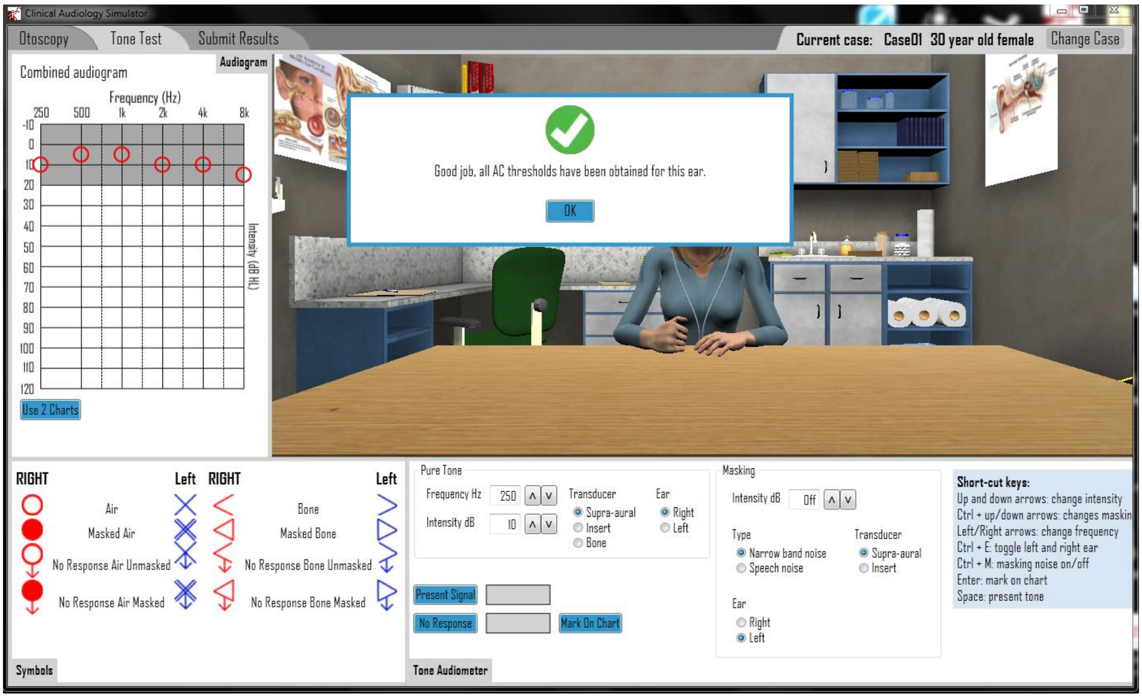



**Figure 3.** *Screen shot of positive reinforcement in response to obtaining all necessary thresholds for one ear*

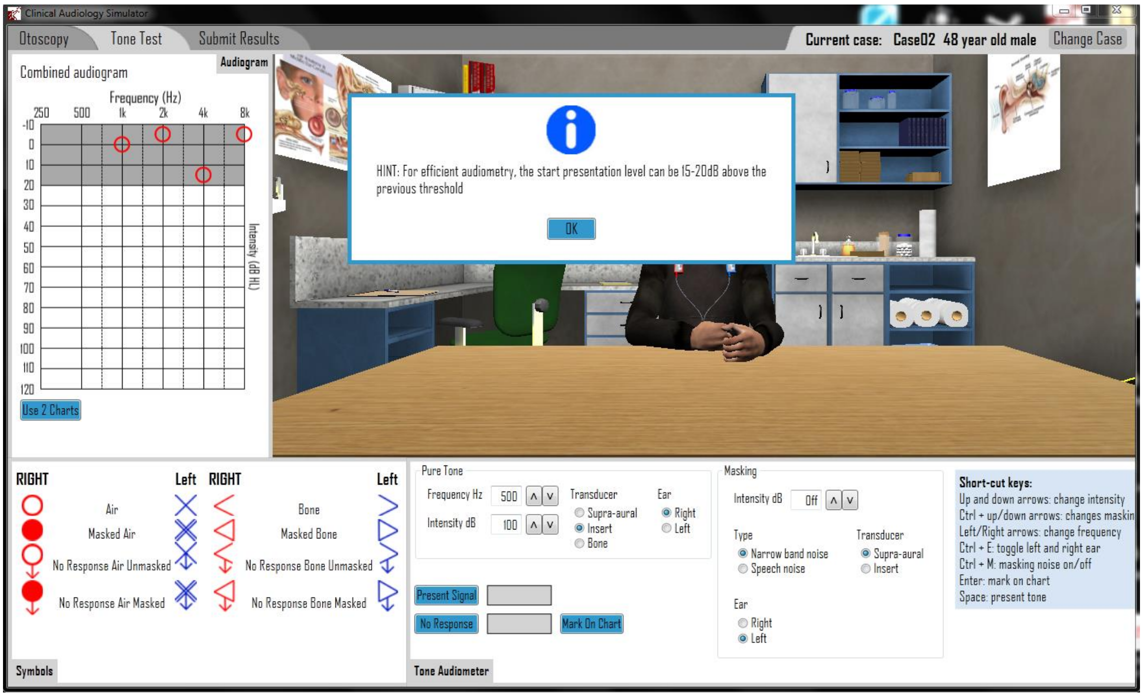



**Figure 4.** ***Screen shot of hint in response to an excessive starting level intensity presented to the VP when testing at 500Hz***

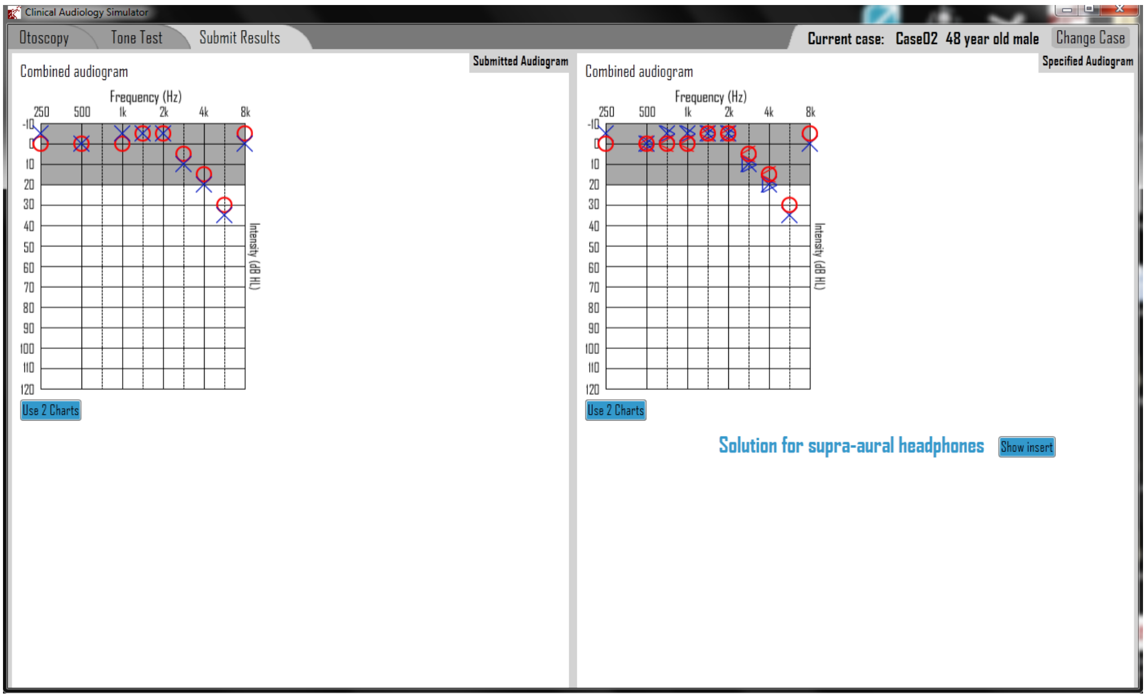



**Figure 5.** ***Screen shot of model answer (on the right) provided after submitting audiogram obtained (on the left) for Case Two***

## 2.2 Participants

All participants were students enrolled in undergraduate Communication Disorders (CMDS) papers at the University of Canterbury, including students in the Bachelor of Speech and Language Pathology with Honours (BSLP(Hons)) course, prospective BSLP(Hons) students completing their 100-level intermediate year and prospective Master of Audiology (MAud) students. The sample consisted of 51 students, 50 females and one male, aged between 18 and 49 years, with a mean age of 23.9 years (SD = 7.80). Twenty-eight of these students were enrolled in the professional degree course of a BSLP(Hons), of which 10 were studying at 200-level, 12 at 300-level, and six at 400-level. Thirteen students were studying at 100-level, undertaking the requirements of the intermediate year of a BSLP(Hons), with the aim gaining entry into the restricted professional course. Ten students were enrolled in other degree courses or had graduated, but expressed interest in applying for the Master of Audiology course (also restricted entry) and were enrolled in at least one CMDS paper. Out of the sample, 17 students reported having some previous audiology experience, such as observing or undergoing a hearing test. The sample was split into two training groups and was counter-balanced according to background and academic year (if in the BSLP(Hons) course), as well as age. Table 1 shows a summary of the descriptive characteristics of the two groups. An independent samples t-test indicated no significant difference in age between the two groups ($t(49) = 0.409$, $p = 0.685$).

### 2.2.1 Ethical Considerations

This project was reviewed and approved by the department of Communication Disorders, the HIT Lab NZ, and the University of Canterbury Human Ethics Committee Low Risk Approval process (see Appendix IV). Undergraduate CMDS students were invited to participate in this study via class visits, emails via the year group course co-ordinators, a student room notice, and word of mouth. Participants volunteered for the study and were recruited on a first come, first served basis. Upon completion of all components of the study, participants were reimbursed for their time with a twenty dollar Westfield shopping voucher, and they entered their names in a prize draw to win either an iPod Touch, a Kindle Touch 3G, or gift vouchers of an equivalent value. All participants signed informed consent forms and were made aware that their participation was entirely voluntary and they could withdraw from the study at any time (also in Appendix IV).

**Table 1. Demographic information for participants in the two training groups**

| | Summative feedback training group | Formative feedback training group |
|---|---|---|
| Number of participants | 26 | 25 |
| Mean age (standard deviation) | 23.5 (7.64) | 24.4 (8.08) |
| Number studying BSLP | 22 | 19 |
| 100-level | 7 | 6 |
| 200-level | 4 | 6 |
| 300-level | 7 | 5 |
| 400-level | 4 | 2 |
| Other* | 4 | 6 |
| Number reporting previous Audiology experience | 10 | 7 |

*studying a degree course other than BSLP or graduated; enrolled in at least one CMDS paper

## 2.3 Procedures

### 2.3.1 Procedure for CAS Training Session

The same procedure was undergone by all participants. Interested students made contact and a CAS training session was scheduled. Training sessions were run with a maximum of five participants at one time according to the availability of computers with the simulator installed. Training sessions were overseen by the MAud student and the PhD student conducting the research. The following protocol was undertaken once a training session was scheduled:

1. **Preparation:** Participants were emailed a basic outline of the clinical protocol for conducting pure-tone audiometry with adult patients (see Appendix V). This document was adapted from the UoC Speech and Hearing Clinic Audiology Protocols and Guidelines (unpublished; 2012), a document guiding student clinical practice at the UoC Speech and Hearing Clinic, and upon which the clinical practice assessment for MAud students is based. Participants were emailed this document the day before coming to the training session, and urged to read through it before attending. In doing so we aimed to familiarise the participant with the PTA vocabulary and procedure. If participants had not read through the information prior to attending the training session they were permitted 15 minutes before commencing training to read the protocol overview. Participants were not permitted to refer to the document at any point during the training session.

2. **Training with the CAS:** Participants were instructed on how to navigate around the CAS before working independently on their allocated version to conduct a hearing test and obtain a virtual audiogram for each of the four virtual patient cases programmed specifically for the training mode used in this study.

    a. Upon opening Case One, a brief passage of case history was displayed (Figure 6), and participants were able to access an otoscopy image and description by clicking the 'Otoscopy' tab in the top left corner of the screen (Figure 7). This information was provided in both versions for all four cases.

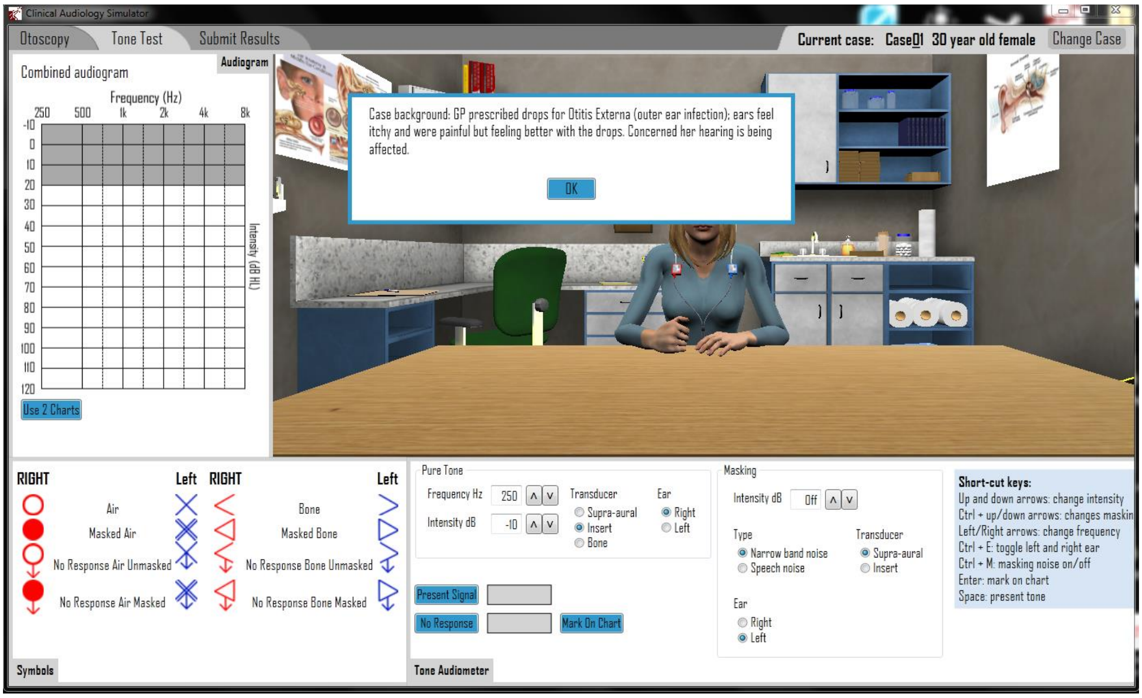



Figure 6. *Screen shot of background history for Case One*

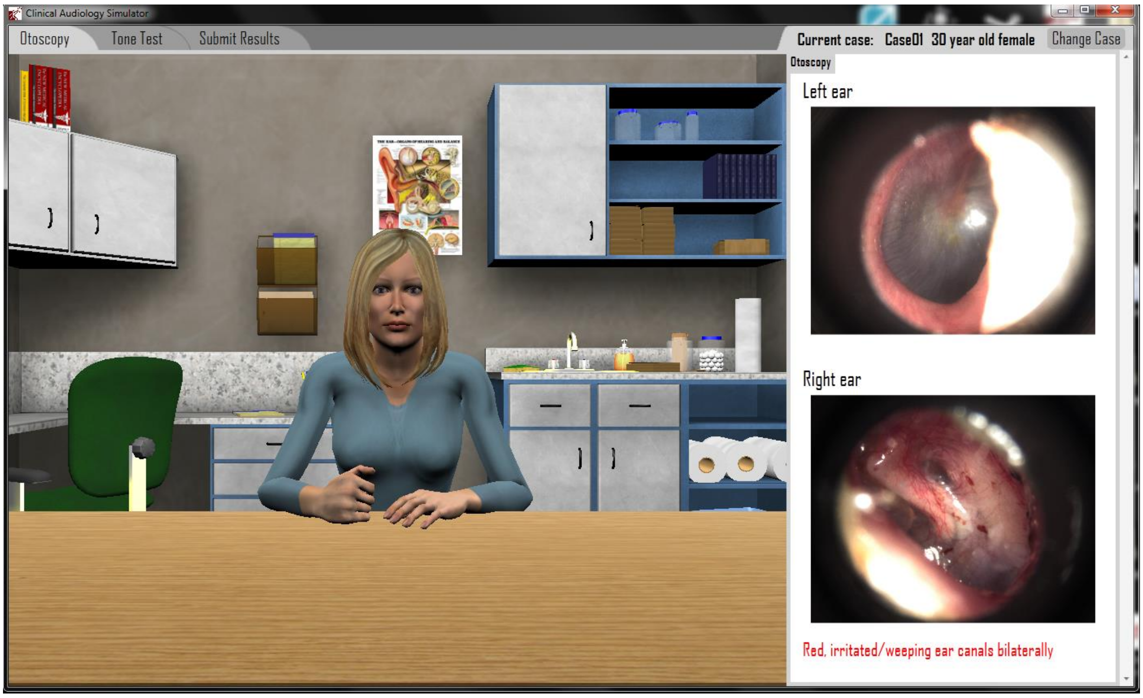



Figure 7. *Screen shot of otoscopy images and descriptions for Case One*

b. Participants then proceeded with PTA.

c. Once a participant had marked all hearing thresholds obtained on the virtual audiogram for Case One, they then submitted their results and the programme revealed the model answer. This happened in both versions of the training mode.

d. Participants then selected Case Two and repeated the procedure, until all four VP audiograms had been submitted, thus completing the training session. All participants worked through the same four cases in the same order (from Case One to Four). There was no time limit for the participants to complete the training session (the training session took under two hours). Participants were offered assistance regarding how to operate the computer programme, but assistance regarding how to conduct PTA was kept to a minimum.

### 2.3.2 Outcome Measurement

A paper-based probes activity was developed and the NASA Task Load Index (NASA-TLX) was utilised in this study for outcome measurement. The development and utilisation of these assessments are described below.

#### 2.3.2.1 Probes activity

The probes (see Appendix VI) were developed for this study in order to examine participants' knowledge of PTA, with the goal of identifying which type of feedback (formative or summative) was most effective in supporting learning of the PTA procedure. This activity served as a type of transfer task to assess a subject's ability to apply their knowledge of the PTA procedure learned from the virtual environment to a paper-based test. Evaluating to what extent a participant can transfer their learned knowledge is thought to be a strong indication of the quality of the learning experience (Bransford et al., 2000).

The probes consisted of a case-based activity done with pen and paper and required participants to answer multi-choice and short answer questions surrounding procedural steps and decision making required when conducting PTA.  Five cases targeting different aspects of the PTA procedure were used. The cases progressed from basic concepts to more difficult concepts. The probes were developed to be consistent with the progression of practice cases within the CAS training mode. It should be noted, however, that all probe cases were novel to the participants. A total of 13 concepts were probed, in the form of 13 multi-choice and 13 short answer questions, with each concept probed with both types of questions. A marking schedule (see Appendix VII) was developed and agreed upon between the researchers and the Clinical Co-ordinator of Audiology, where 22 of the questions were allocated 1 point, and 4 allocated 2 points, totalling a maximum possible score of 30 points.

#### 2.3.2.2 NASA Task Load Index (NASA-TLX)

The NASA-TLX is a subjective workload assessment. Developed by the Human Performance Group at NASA's Ames Research Center in the late 1980’s, the NASA-TLX provides a way to assess subjective workload of those working with various human-machine systems, such as simulators. The assessment consists of a two-part evaluation procedure from which an overall workload score is derived. Six categories (mental, physical and temporal demands of the task, and own performance, effort and frustration during the task) are individually rated on a likert scale according to how important each was in creating workload during the task. The categories are then given a weighting based on a series of pairwise comparisons, where the category that was perceived to contribute most to the workload of the task is identified. The overall workload score for each subject is calculated by multiplying each rating by the weight given to that category by that subject, and then the sum of the weighted ratings is divided by 15. The higher the overall workload score, the higher the perceived workload[2].

---

[2] Further information including the resources used for conducting the pencil and paper version of the NASA-TLX is available from:
http://humansystems.arc.nasa.gov/groups/TLX/downloads/TLX_pappen_manual.pdf

### 2.3.3 Procedure for Outcome Measurement

The procedure for conducting outcome measurement during the study is described below. The overall study design is summarised in Figure 8.

1. **Subjective assessment of workload:** After completing the training session with the CAS, participants carried out the paper version of the NASA-TLX to obtain a subjective assessment of workload whilst undertaking training with the simulator.

2. **Probes of pure-tone audiometry knowledge:** Within 24 hours of undergoing the CAS training session, participants returned to the training venue to conduct the follow-up session. Participants completed the probes activity, working independently with no time limit (the exercise took under an hour). The probes activity was administered by the MAud student and/or the PhD student conducting this research.

3. **Repeat NASA-TLX:** After finishing the probes activity, participants completed the NASA-TLX a second time to obtain a subjective assessment of workload whilst undertaking the probes activity.

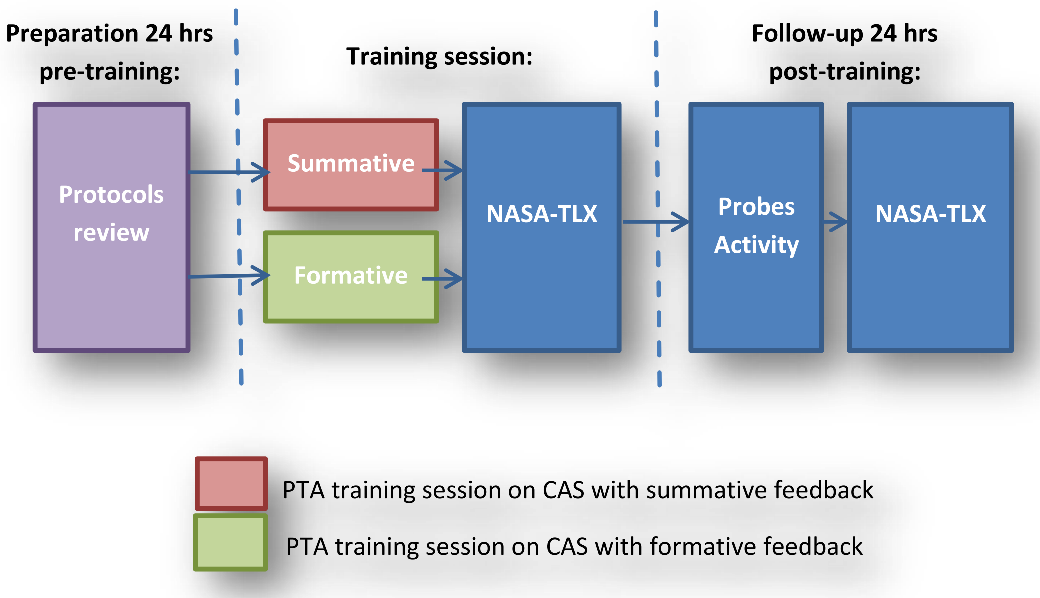



Figure 8. *Study design diagram*

### 2.3.4 Other Information Gathering

Participants completed a Participant Information Form (Appendix VIII) to obtain demographic information in order to identify any variation between the two feedback groups, and a Usability Questionnaire (Appendix IX) for participants to offer feedback for refinement of the software design.

Timing information was also collected**.** The duration of participants' individual training sessions with the CAS was logged and extracted from the software. The time it took for each participant to complete each case was recorded in minutes and seconds.

### 2.3.5 Statistical Methods

Data for each participant was collected in a Microsoft Excel spread-sheet. Between-group means from the probes activity and NASA-TLX were compared using independent sample t-tests, calculated in The Statistical Package for the Social Sciences (SPSS). A significance value of $p \leq 0.05$ was used throughout. Within-group regression analysis of timing data was conducted using Microsoft Excel.

# 3 RESULTS

## 3.1 Probes

Probes were administered in order to evaluate learning of the PTA procedure. Participants (n = 51) completed the probes within 24 hours of undertaking training with the CAS whilst receiving either summative or formative feedback. An independent samples t-test was conducted in order to determine if the two training groups differed in accuracy in the probes. It was found that participants (n = 26) who received summative feedback achieved a mean score of 14.63 (SD = 4.20). In contrast, participants (n = 25) who received formative feedback achieved a mean score of 20.88 (SD = 3.54; Fig. 9). The difference was significant with those receiving formative feedback achieving a higher score on the probes activity ($t(49) = 5.73$, $p < 0.0005$, $d = 1.64$). When referring to effect size magnitude guidelines this is a large effect (Cohen, 1988, 1992).

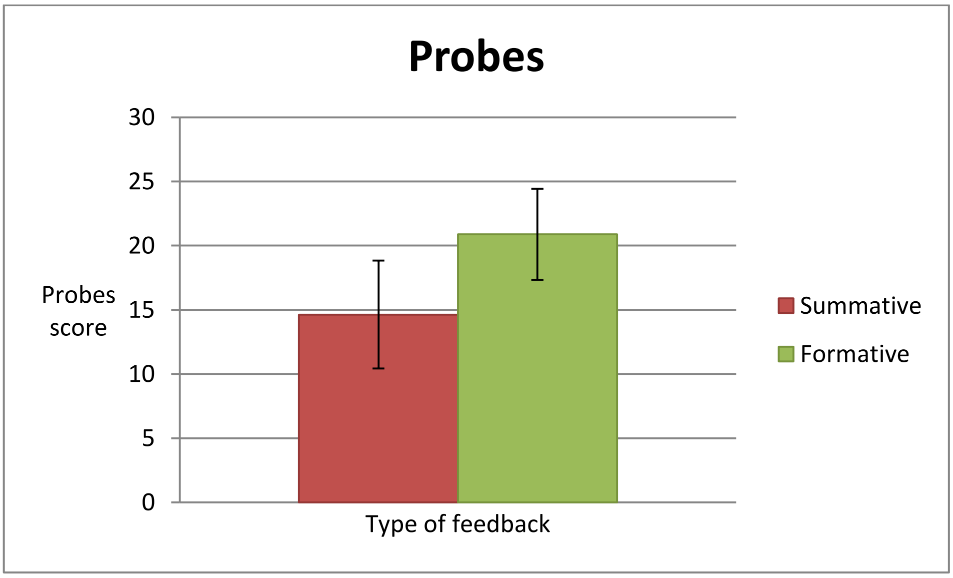



**Figure 9.** ***Probes activity mean scores according to type of feedback received during PTA training. Error bars show standard deviation***

## 3.2 NASA-TLX

All participants completed the NASA-TLX twice, once following PTA training with the CAS, and again following the probes activity, in order to evaluate perceived workload during these two tasks. An independent samples t-test was conducted in order to determine if perceived workload differed between the feedback groups during PTA training. Participants receiving summative feedback reported a mean overall workload score of 47.77 (SD = 14.10), whereas participants receiving formative feedback reported a mean overall workload score of 60.19 (SD = 13.54; Fig. 10). The difference was significant with those receiving formative feedback reporting a higher overall workload score ($t(49) = 3.21$, $p = 0.002$, $d = 0.92$). This is a large effect (Cohen 1988; Cohen, 1992). A further independent samples t-test was conducted in order to determine if perceived workload differed between the groups during the probes activity. Those receiving summative feedback reported a mean overall workload score of 61.56 (SD = 11.11), while those receiving formative feedback reported a mean overall workload score of 54.48 (SD = 13.0; Fig. 10). Again, the difference was significant, this time with those receiving summative feedback reporting a higher overall workload score ($t(49) = 2.10$, $p = 0.04$, $d = 0.6$). This is a medium effect (Cohen 1988; Cohen, 1992).

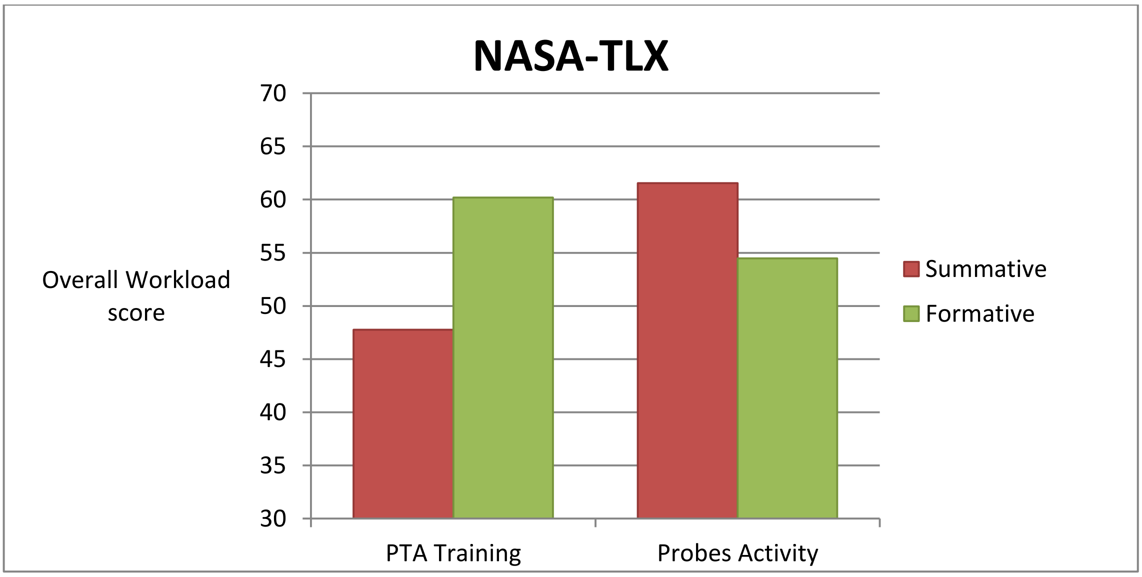



**Figure 10.** *Mean overall workload scores during PTA training and during probes activity according to type of feedback received*

## 3.3 Timing Information

The amount of time participants spent using the simulator during the training session was analysed to determine the effect this had on learning the PTA procedure (assessed with probes). It may be proposed that the longer a participant spent training, the better their score, regardless of the type of feedback they received. Participants worked through the four cases without any time limitation. The time it took for each participant to complete each case was recorded by the computer. Due to a system error, accurate timing information was only available for 21 participants from the summative feedback training condition, and 13 participants from the formative feedback training condition. Timing results for participants according to the type of feedback they received are reported in Table 2. A within-subjects regression analysis was conducted for the participants in each training group, comparing total time taken to complete the training session and probes activity scores. Scatter plots of these analyses are shown in Figures 11 and 12. No correlation between time and probes score was found for those who received summative feedback whilst training ($R^2$ = 0.08), nor those who received formative feedback ($R^2$ = 0.07), therefore, there was no relationship between spending time on the simulator and outcome on the probes.

**Table 2. Mean time and standard deviation (hours:minutes:seconds) taken to complete each case and entire training session by feedback condition**

| | Summative feedback | Formative feedback |
|---|---|---|
| **Mean time (SD) – Case 1** | 11:51 (4:34) | 10:43 (3:53) |
| **Case 2** | 12:43 (7:16) | 11:51 (2:58) |
| **Case 3** | 11:25 (5:02) | 27:23 (7:23) |
| **Case 4** | 16:04 (8:50) | 23:48 (8:20) |
| **Total mean time (SD)** | 49:53 (17:16) | 1:08:30 (25:47) |

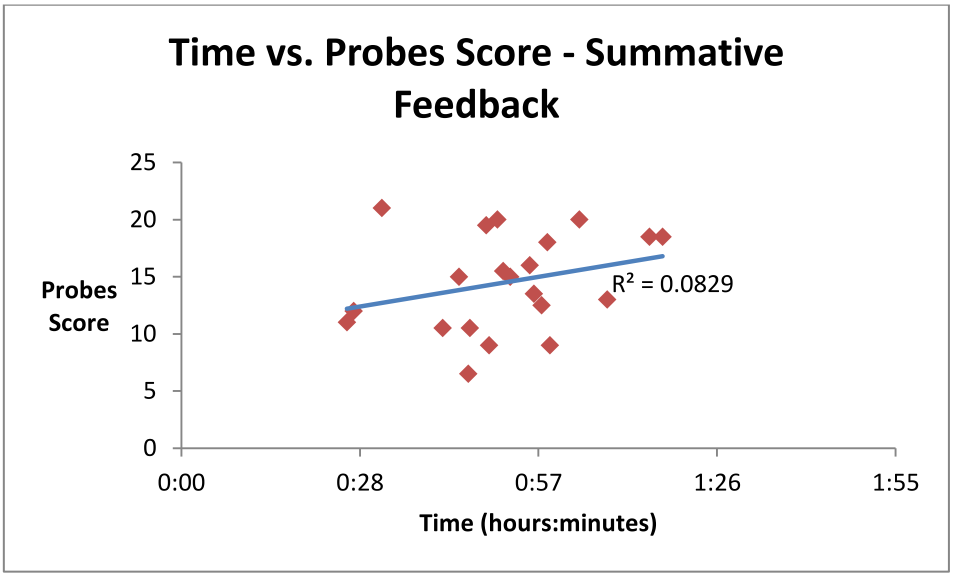



**Figure 11.** *Total time taken to complete PTA training compared to probes score of participants who received summative feedback*

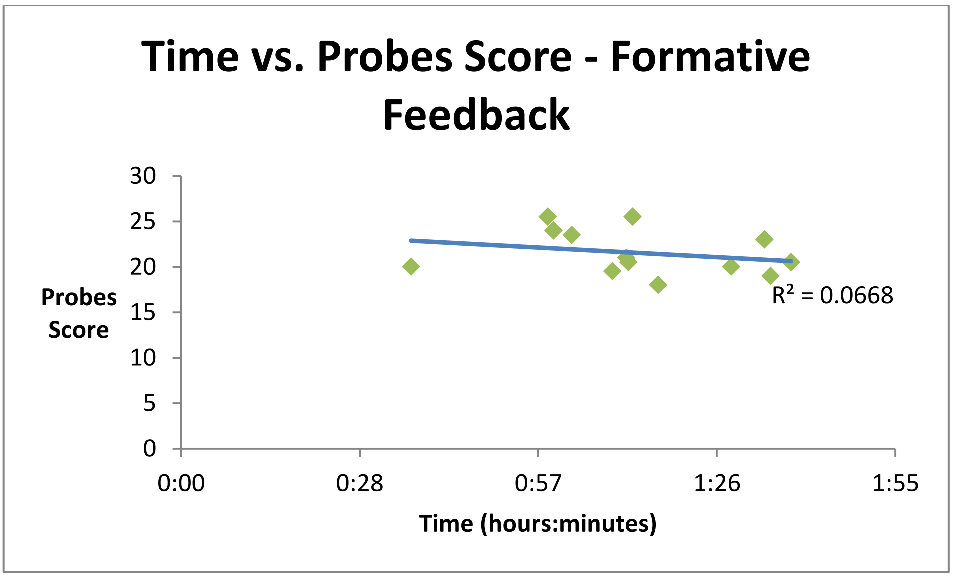



**Figure 12.** *Total time taken to complete PTA training compared to probes score of participants who received formative feedback*

# 4 DISCUSSION

## 4.1 Aim One

The first research objective was to determine whether participants would demonstrate better knowledge of the PTA procedure after receiving formative feedback. It was hypothesised that the formative feedback condition would lead to a higher degree of knowledge acquisition and ability to apply this knowledge than learners who received summative feedback.

The formative feedback training version dispensed elaborative feedback immediately in response to mistakes, only allowing the subject to continue testing once they had performed the procedure correctly. Also, response-contingent hints and positive reinforcement were provided throughout the hearing assessment of each VP. Subjects were also able to see a model answer upon completion of a worked case. The results confirmed the hypothesis by showing that on average, participants who trained on the CAS with formative feedback scored higher in the probes activity than those who received summative feedback.

The summative feedback condition involved implicit verification feedback only, in the form of revealing a model answer once a worked case was submitted. Participants were able to compare their audiogram obtained with the model answer, thereby deducing whether or not they had obtained the correct thresholds of hearing. There was no indication as to whether they had gone about the PTA procedure correctly, or why their audiogram may have been different to the model answer. Participants receiving this type of feedback were clearly outperformed in the probes activity by those who received formative feedback, indicating the summative condition was less effective in supporting learning of the PTA procedure in novice learners.

This finding is supported by the literature surrounding feedback in traditional educational settings. Generally, more elaborative feedback that includes aspects of both verification and elaboration is reported to be more effective in supporting learning than verification feedback alone. In their review, Kulhavy and Stock (1989) acknowledged this was an assumption of information-processing, being that error correction would more likely result from more information provided in the feedback message. Whilst studies included in their review failed to support this assumption, a meta-analysis conducted a short time later by Bangert-Drownes et al. (1991) observed the pattern that more elaborative feedback strategies were more effective than simple verification feedback where outcome measures included the application of information, requiring understanding of the information acquired with the assistance of feedback. Bangert-Drownes et al. (1991) indicated that elaboration feedback was important for the conceptualisation of information, ability to make inferences and apply rules. This could explain why training with the formative feedback version of the CAS, which included more elaborative feedback, was more effective in facilitating performance when assessing understanding of a clinical procedure with a transfer task.

Other studies within the realm of computer-based instruction utilised transfer tasks as outcome measures when determining the effect of elaborative feedback on learning. Studies by Moreno (2004), and, most recently, Butler et al. (2013), are examples where students who received elaborative feedback during computer-based training were found to produce higher scores on transfer tasks when compared to verification feedback. Butler et al. (2013) argue that the most important aspect of any feedback procedure in terms of transfer of knowledge is the content of the feedback message alone. In considering elaborative feedback as an isolated variable, they found that the additional information fostered better comprehension of the material, moving the learner from superficial factual knowledge to a more complex understanding of the concepts.

It must be acknowledged that elaborative feedback was not observed in isolation in this study. Perhaps of more relevance is literature in the line of Moreno's (2004) research that contributes to developing design principles for educationally effective interactive environments, including simulations (e.g. Moreno & Mayer, 2007; Plaas, et al. 2009). These guidelines include feedback among a number of variables to be considered for such systems to enhance learning. Whilst not specifically related to computer learning environments, but still relevant when considering feedback design for such systems, Shute's (2008) review and formative feedback guidelines also follow this reasoning, acknowledging that feedback and learning are entwined, multifaceted and that many variables impact on both. The design of the CAS formative feedback training mode, whilst including more elaborative feedback, also included other features such as increased frequency and immediacy of feedback. These other features are likely to have also supported subjects' abilities to apply aspects of the PTA procedure to a new context in addition to a more elaborative type of feedback.

One such feature of the formative feedback training mode was to offer increased immediacy of feedback timing. As soon as a subject performed an incorrect action, a feedback message popped up with information reiterating the protocol relevant at that point, from which subjects could extract the correct answer, repeat the action correctly and move on in conducting the hearing assessment. Meta-analysis by Kulik and Kulik (1988) found that immediate feedback was more beneficial when tasks demanded more complex conceptual learning. Immediate feedback was also considered by Shute (2008) as a favourable condition for difficult tasks and for novice learners. Thus, the formative feedback training mode took into consideration the complexity of the task and knowledge level of the novice learners more appropriately.

The summative feedback condition had a more delayed form of feedback timing, where the model answer was only revealed upon completion of a case. Shute (2008) suggested that delayed feedback may promote transfer of learning, but this was not observed in the current study. Participants in the summative feedback group performed worse in the transfer task than those who received more immediate feedback. However, timing of feedback as a variable was not considered in isolation, thus no further information can be added to the research from which Shute (2008) drew that conclusion. Feedback type is possibly a more powerful variable than feedback timing, but in order to investigate this proposal, further research is required. It would be interesting to study two versions of the CAS formative feedback training mode where feedback timing (immediate or delayed) was manipulated to see whether this further affected transfer of knowledge, or whether it is a result of the quality of the feedback message alone as Butler et al. (2013) propose. If so, the supposed result would be equivalent performance on the transfer task regardless of timing condition.

Despite clear learning benefits of the formative feedback training mode, it is important to consider how the participants themselves felt about using the CAS. Upon completion of the training session, participants had the opportunity to write about positive and negative aspects of their experience with their respective version, in order to provide information for improving the simulator, and to consider face validity and external validity of the CAS.

## 4.2 Participant Feedback

For many of the participants who received summative feedback whilst training, the model answer was acknowledged as a positive aspect of the CAS. Furthermore, being able to practice hearing testing in the virtual environment was an exciting prospect for most (e.g. "was able to practice audiometry as I've never had the chance previously"). Further comments recognised the benefits and potential uses of the software (e.g. "portable tool, you could practice anywhere" or "easily distributed/made available to large groups so there could be a mass training initiative"). Comments such as: "it helped to facilitate learning without using a real patient", "extremely helpful in reinforcing how to do a hearing test" and "reinforces concepts learned in class" showed that the learning power of the summative feedback version should not be undervalued. This type of feedback would be better suited to more experienced users, where delayed and/or verification feedback is considered to be most beneficial (Shute, 2008).

The negative aspects noted by those in the summative feedback condition predominantly surrounded the lack of guidance. General comments included "no guidelines", "there wasn't much guidance as to whether you were testing the right thing" and "I didn't know if what I was doing was right". More specifically, participants felt they needed help when decision making, reflected in comments such as: "No feedback or instructions on when different headphones were necessary", "No feedback on decisions about how many frequencies to test". Also, the summative nature of the feedback was perceived by some as a negative, in that "once results were submitted they couldn't be edited" or "can't have multiple attempts after seeing the answer", and the lack of elaboration was also seen negatively, e.g. “No feedback on why I got bits [of the audiogram] wrong”. Many wrote their negative aspects as suggestions of how to improve the CAS. These have been compiled in Table 3, which also makes a note of which of these suggestions were addressed by the formative feedback training mode, that was developed prior to when participant feedback was obtained.

Those in the formative feedback condition positively commented about the CAS as a training tool, noting that practicing with the virtual patients was “better than learning from a textbook” and a “useful way to learn how to do a hearing test”. Generally, participants felt it “didn’t take too long”, was “very easy to master [in terms of navigating the software]” and “helped me understand the rules behind hearing testing”. Encouragingly, some participants noted it was “really practical, would be great for training Audiologists to practice with” and an “excellent programme, wish they had this for the 242 course” in reference to the undergraduate paper, CMDS 242: Introduction to Audiology. Positive comments about the feedback messages included “helpful when I am incorrect/helps me understand”, “corrections are good in preventing you going too far in the wrong direction”, and that the “tips guide you but you still have to figure the answer out”, indicating that some participants felt being “encouraged to make mistakes for learning” was constructive. More specifically, the feedback messages were seen to be helpful in indicating completeness of the procedure, e.g. “won’t let you finish without completing all thresholds” or “let’s you know when you had enough to move onto the next case”. Also, for more technical aspects of the PTA procedure, participants commented positively that the CAS could provide cues for the necessary calculations, such as initial masking level. Interestingly, one subject commented that it was “good to see answer at the end when you submit” suggesting that revealing the model answer was still valuable in the formative feedback condition.

Table 3. Participant suggestions for improvement of the CAS

| Suggestion | Addressed by FF version? |
|---|---|
| Give clear prompts as to which headphones to use | ✓ |
| Further explain (within the programme) the use of bone conduction | ✓ |
| Give prompts throughout if you were doing things wrong | ✓ |
| Give instructions & explanations of tests | ✘ |
| Provide hints for beginners | ✓ |
| Be able to go back | ✘ |
| Have an information button for reminders of what each thing is, e.g. what bone conduction threshold to use | ✘ |
| Provide (written) instructions during use | ✓ |
| Provide feedback if thresholds were incorrect and why | ✓ |
| Provide information regarding the diagnosis | ✘ |
| Provide information as to why your final answers were different | **N/A*** |

FF – Formative Feedback

*the formative feedback version doesn't allow the user to submit an incorrect/incomplete audiogram thus it is not possible (in theory) to get a different audiogram to the model answer.

Negative comments about their experience suggest the formative feedback version was very demanding for these participants. Numerous negative comments surrounded the complexity of the feedback messages, indicating some explanations were too long, or some instructions were not that clear, e.g. "didn't always know what particular area instructions were referring to" and "feedback wasn't specific enough". In particular, "masking prompts were difficult to understand" and "masking concept was difficult to understand" were comments that suggested this more complex aspect of PTA was not well supported by the formative feedback for some. Another negative aspect was the difficulty in dealing with errors that many subjects experienced, summarised neatly with the following comment: "difficult to recover following a mistake - if you muck up it is sometimes difficult to fix". Subjects often found themselves in a feedback loop as a result of not being able to retrace their steps to the point before they had gone wrong, e.g. "when threshold seeking, if a window pops up, it doesn't tell you where you were". The software was not designed to perform this action and subjects just had to remember and reset the audiometer settings themselves, meaning they had to be very mindful of their actions throughout the training session. This proved to be a source of frustration for some, e.g. "I felt frustrated when I couldn't figure it out" and that it was "frustrating when I tried to do something and I didn't know why it wouldn't let me".

Some aspects of the formative feedback condition that were seen as positives were considered negatives by others. Comments noted that the CAS "doesn't tell you what to do next" and that it was "frustrating to have to get it wrong before you received feedback", which indicated some learners disliked having to make errors, and desired a more explicit and directive experience, which Shute (2008) suggested can be more constructive with novice students. In addition, the strategy where information was only available when a mistake was made was perceived as being less constructive for some. Interestingly, one subject noted that there was "no overall feedback, e.g. what you did well/poorly at" indicating a summative evaluation of performance may have been helpful for this person. Again, participants listed suggestions of improvements by way of providing negative critique, and these have been taken into account and discussed in Recommendations, section 4.6 below.

It is concerning that many of the participants in the formative feedback condition felt frustrated and confused when training with the CAS. Despite this, the group performed better on the probes activity, indicating a more enhanced understanding of the PTA procedure than those who received summative feedback. Discussion of the second aim of this study provides more insight as to why this may have occurred.

## 4.3 Aim Two

The second research objective was to determine how cognitive load would differ during training and during assessment between the feedback conditions. Cognitive load is the demands placed on working memory, involved in central information processing, during activities that facilitate learning. According to De Jong (2009), instructional designs should consider the prior knowledge of the learner and avoid non-essential or confusing information and stimulate processes leading to deep knowledge acquisition. The NASA-TLX is a commonly-used assessment for cognitive load. Considered a multi-dimensional subjective measure of workload, the NASA-TLX takes into consideration the processing resources of the subject and the demands of the task. Hart and Staveland (1988) describe workload as being the perceived relationship between the mental processing abilities of a learner and the processing requirements of a task. Participants in this study completed the NASA-TLX after training with the CAS and after undertaking the probes activity in order to obtain a subjective evaluation of cognitive load during the respective tasks. The NASA-TLX provided an overall workload score, where increased perceived workload was reflected by a higher score.

The hypothesis was that those receiving formative feedback would experience a higher perceived workload during PTA training with the CAS, but this would reduce during assessment. Those who received summative feedback would experience the opposite. The results confirmed the hypothesis, showing that whilst training with the CAS, participants who received formative feedback reported an average overall workload score that was significantly higher than the group who received summative feedback. When undertaking the probes activity, the formative feedback group experienced significantly lower scores on average than the summative feedback group. The formative feedback group experienced a 30% decrease in perceived workload from training to assessment, whereas the summative feedback group experienced a 10% increase.

This result suggests cognitive load was higher for those in the formative feedback condition during training with the CAS, indicating that the differences in the design of the formative feedback training mode created more demand. When referring to the literature, this result seems counter-productive, as much research is aimed at designing instructional systems that reduce cognitive load for optimising knowledge acquisition (eg: Paas & Merrienboer, 1994; Moreno, Mayer & Lester, 2001; Moreno, 2004). Design strategies were employed to control for extraneous load, such as presenting training material in a simple-to-complex manner, integrating elaborative feedback as an information source, and allowing this information needed to complete the task to be reviewed without any time constraints (Merrienboer & Sweller, 2005; Plass et al., 2009; Jong, 2009). However, other design features intended to support learning may have been detrimental to cognitive load. For example, some of the feedback messages were not considered to be well implemented, consisting of too much or too complex information, especially for masking which is a more advanced technique in the PTA procedure (carrying high intrinsic demand). Fiorella, Vogel-Walcutt & Schatz (2011) reported that implementing real-time feedback in such environments can be challenging, as feedback may disrupt learners and distract from the task, with the two features competing for cognitive resources. If processing demands when learning from interactive environments exceed the processing capacity of the cognitive system, cognitive overload occurs (Moreno & Mayer, 2007), compromising the ability to learn, resulting in poor performance on subsequent assessment or situations requiring knowledge application.

Despite reporting a higher workload than the summative feedback group, those in the formative feedback condition performed better in the assessment, indicating an increased knowledge gain. This suggests that subjects receiving formative feedback were not cognitively overloaded, and that the difference in demand created by the formative feedback version may have in fact optimised learning. Paas and van Merrienboer's (1994) study found that students learning geometry procedures yielded acceptable levels of cognitive load and showed superior performance in transfer tasks when they studied problems designed to have a low extrinsic load yet high germane load. Additionally, Moreno (2004) found that students with low prior-knowledge learned best when guided by elaborative feedback, and proposed that the more elaborative feedback style reduced extraneous cognitive load, relieving the effort of searching for a plausible answer to their problem. This resulted in freeing capacity for increased germane load, promoting learning. Germane load is induced by learners' efforts to process and comprehend the material for deeper conceptualisation and understanding. It is proposed that whilst recipients of formative feedback reported higher cognitive load during training, perhaps extrinsic sources of load were for the most part appropriately supported, and germane load was increased. This promoted a better understanding of PTA and an increased ability among these subjects to apply this information to a new context. This is an assumption, as a further disadvantage of the workload assessment (and the majority of other cognitive load measures as discussed by De Jong, 2009) was that the overall workload score did not indicate which demands contributing to cognitive load (intrinsic, extrinsic or germane) impacted on subjects the most.

In contrast, subjects who received summative feedback reported a lower overall workload score during training, but did not perform as well in a transfer of knowledge task, suggesting they did not undergo the process of learning as thoroughly as the formative feedback group. As there was no guidance *during* the task of conducting a hearing test, the attention of learners may not have been appropriately directed to processes thought to be relevant to learning, such as building understanding and generalising meaning (Sweller et al., 1998). This suggests subjects in the summative condition experienced sub-optimal levels of germane load. Also, extrinsic cognitive load was reduced due to the absence of pop-up feedback messages in comparison to the formative feedback condition. The only factor contributing to subjects' perceived workload in the summative condition was the intrinsic complexity of the task. Whilst this likely resulted in an overall workload score that was lower than the formative feedback group, participants' comments indicated they did not feel appropriately supported during training, and achieved lower scores in the probes activity, suggesting the summative feedback training mode of the CAS did not place appropriate demands on cognitive load, and learning suffered as a result.

Furthermore, the summative feedback condition did not provide subjects with the appropriate information to complete the assessment task, meaning they were relying solely on their prior knowledge and intuition to answer the probes. As a result, this task was perceived by these subjects to be more demanding than the training session, as it commanded answers about concepts they may have only read about and may not have mindfully practiced or experienced when conducting the PTA training. That is, the task likely carried a far greater intrinsic cognitive load due its greater complexity for these subjects when compared to the training session. For the formative feedback group, the assessment task was perceived to be much less demanding than the training session. These subjects had their attention drawn to important concepts of the PTA procedure by the pop-up feedback messages and mindfully executed these concepts in order to progress through the PTA training. Similar concepts were probed by the assessment task, which likely resulted in reducing the intrinsic load of the task. Also, extrinsic load was possibly not as high for the formative feedback group as it was during training due to the absence of the confronting feedback messages during the probes activity. Again, these are assumptions since the workload assessment did not consider cognitive load by the different contributors as defined by cognitive load theory.

## 4.4 Timing

This study also investigated a secondary question, namely whether time spent training with the CAS had an effect on learning, irrespective of the feedback variable. On average, the formative feedback group spent more time training on the CAS (25% longer than the summative feedback group), and this group performed better than the summative feedback group on the probes activity. Results of the regression analysis, however, revealed no relationship between participants' probe scores and the time taken to complete the training session. This suggests that more time spent working through the cases did not lead to a better understanding of the PTA procedure, but reinforced a theory of deliberate practice, where the quality of the experience has more of an impact than quantity alone when gaining in clinical skill proficiency (Duvivier et al., 2011). The nature of the formative feedback forced participants to spend more time reading, interpreting and applying the feedback messages, which overall meant more time was taken to work through the four cases. This somewhat reduced efficiency, a commonly reported criticism of formative feedback (e.g. Mason & Bruning, 2001), is of minimal concern considering the demonstrated learning benefits, where subjects gained a deeper understanding of the PTA procedure, reflected by higher scores in the transfer of knowledge task.

## 4.5 Limitations

In addition to the strengths of this study, it is important to consider the limitations in the methodology and the sample of individuals involved when interpreting the findings described above.

### 4.5.1 Sample

Basic knowledge of PTA and audiogram interpretation is important for Speech-Language Pathologists. The clinical service of screening individuals for hearing loss or middle ear pathology using conventional pure-tone air conduction methods (including otoscopic inspection), is within a Speech-Language Pathologist's scope of practice according to the American Speech-Language-Hearing Association (2007), as is providing services to individuals with hearing loss. Furthermore, collaborating with other professionals, such as audiologists, is an important element of the role, hence the benefit to prospective Speech Language Pathologists of gaining further audiological experience, and the inclusion of this body of students in the current research.

Undergraduate CMDS students obtain foundation knowledge of PTA via a 200-level course, CMDS 242: Introduction to Audiology (in 2012 this course was offered in the first semester, February to June). Other 100-level intermediate year CMDS courses also discuss aural physiology and functioning, and the importance of hearing within the realm of speech and language therapy. These courses are compulsory to those students enrolled in the BSLP(Hons) programme, and are recommended for those hoping to gain entry to the MAud course. Despite gaining a theoretical basis via these courses, students have limited practice and experience in conducting PTA, with few opportunities to consolidate this knowledge before graduating. Unlike post-graduate students enrolled in the MAud course, those enrolled in undergraduate CMDS courses are not consistently exposed to or practicing audiological assessments, thus the nuisance variable of concurrent experience or exposure to PTA alongside this study was reduced, isolating the effects of the feedback conditions in the CAS.

It was also anticipated that by recruiting students from the undergraduate CMDS courses that a large number of participants would be available for this study. Annually 30-40 students are enrolled in each of the three professional years of the BSLP(Hons) course, with a larger number of students considering applying for the BSLP(Hons) course at the intermediate level. Also contributing to the pool of available students are those that are following other degree courses, or those that have graduated, who take CMDS papers for interest's sake. It was estimated that upwards of 60 participants would be recruited from a pool of over 150 students. Despite active recruitment over six months and study sessions running over four months, a relatively small sample size could be recruited for this study.

It is likely that the intensive nature of the BSLP(Hons) programme meant students were hesitant to take time out from their very busy schedules to participate in this study. Furthermore, students at the 300 and 400 levels were required to undertake placements outside of the Canterbury area, meaning a number of students did not express interest in participating. For students that did volunteer, it was not always possible to arrange sessions that suited everyone's schedules, despite considerable efforts to do so. Some became unable to participate due to illness, a disadvantage of running study sessions throughout the winter months. Offering a different incentive to participate, such as a small percentage of course credit, may have boosted the sample size. Certainly, more undergraduate students were available than if MAud students alone were involved in the study, which number just 12 on average for each of the two years of the programme.

Participant self-selection is a potential limitation of this study which may have biased the sample and results in some way. That is, students who were particularly interested in Audiology or studying Audiology may have felt that participating in this study was a way of gaining experience and improving their chances of being accepted into the MAud course. Alternatively, those who were not at all interested in Audiology may have chosen not to be involved. As a result the sample may be biased towards individuals who had more motivation to practice PTA. This may be a positive bias considering the potential CAS users, who will most likely be following the MAud degree course, or in some way be interested in audiometric assessment. That is, users will most likely have some motivation to practice PTA. However, these findings may not necessarily translate to CAS users who are not characteristic of the sample, that is, not university students, or those beyond novice level.

Lastly, it must be acknowledged that only one male participated in this research. Having just one male participant reflects the very low rate of male enrolment typical to undergraduate CMDS courses at UoC, and whilst it was hoped more male students would participate, there were few from which to draw a sample from in the first place. Nevertheless, this means that the findings of this research occurred in a predominantly female cohort, and may not extend to male learners. It is likely men and women respond to feedback differently, as other variables that can impact on the effectiveness of feedback, such as motivation and self-esteem, have been proven to manifest quite differently between genders (Cross & Madson, 1997). Regarding feedback in computer programmes, Djamasbi and Loiacono (2008) investigated how computer-based feedback was used and affected mood differently in men and women. Their software delivered summative, verification feedback within a computerised interactive decision support system, which was shown to affect the overall mood of females more negatively than males, and improved the decision accuracy of the female users more so than the males. More research is required to examine gender differences in investigations such as the current study.

### 4.5.2 Methodology

Learning outcomes were measured by a probes activity created for this study. Probes are a commonly used informal assessment tool for gaining an indication of learning, typically designed by educators for gaining useful information about whether instruction has been understood or requires modifying in some way (Cotton, 1988; Keeley, Eberle & Farrin, 2005). This particular probes activity consisted of a short quiz with multi-choice and short answer questions to probe subjects' knowledge of the PTA procedure. Note not all the rules of the procedure were tested; selected basic to more complex concepts were probed in order to be representational of the entire procedure. The activity was case-based and required subjects to apply their experience with the CAS to the new pen and paper context. Whilst the outcome provided a quantitative way to compare the performance of the groups and reflected a difference in knowledge gains, it is unknown whether this was a particularly valid or reliable assessment tool for probing students' understanding as well as their ability to transfer their knowledge to novel contexts. A more valid way of assessing the subjects' abilities to apply their knowledge post-training with the CAS would have been via a practical assessment with the typical clinic set-up, where subjects would have conducted PTA with a standardised patient. Such practical assessments are part of the traditional assessment of MAud students during their training at UoC. However, this would have been very time-consuming and impractical to conduct with the number of subjects required for a powerful study, hence the alternative probes activity.

Cognitive load was measured using a subjective multi-dimensional assessment of workload, namely, the NASA-TLX. This assessment is widely used and is considered to be an accurate and valid measurement of workload (Miller, 2011). This tool does, however, have limitations, which affected the strength of conclusions drawn about the effects of feedback on cognitive load in this study. One limitation of the NASA-TLX is that it was unable to indicate whether workload was so high during the task that learning became impaired. That is, there was no identified level or score that, once breached, indicated cognitive overload. This is a reported flaw of many cognitive load measures, and further research is required to develop measures that quantify cognitive overload (Jong, 2009). As opposed to *general* overload, it is possible that subjects experienced *momentary* cognitive overload; however, a further disadvantage of the workload assessment - as a summative assessment of overall load throughout the task - is that it did not provide information about instantaneous cognitive overload (Jong, 2009). Therefore it was unhelpful in identifying if there were particular points of high-pressure during training with the CAS which could be relieved with additional support. In order to identify instantaneous cognitive load a different technique, such as a physiological measure that records responses throughout the task, may have been more useful (Miller, 2001). A further disadvantage of the workload assessment (and the majority of other cognitive load measures as discussed by de Jong, 2009) was that the overall workload score did not indicate which demands contributing to cognitive load (intrinsic, extrinsic or germane) impacted on subjects the most. Research is underway to develop more sensitive measures of the three mental workload categories (e.g. Galy, Cariou & Mélan, 2012).

## 4.6 Recommendations

### 4.6.1 Changes to the CAS Formative Feedback Training Mode

Utilising the formative feedback training mode of the CAS in this study served to not only obtain answers to the research objectives, but to also test the practicalities of the implementation of the feedback messages. As a result, some minor design changes have been identified. It is recommended that modifications be made to increase the utility of the information within the message, as a number of participants commented they were confused by some of the feedback messages that were too long or complex. Suggested options to make the information more readable may be to use bullet points within the pop-up message so as to chunk the information into more digestible units. Alternatively, a single, lengthy message could be broken into several messages and presented in a more specific, step-wise fashion. This may be particularly helpful for the message about the Plateau Method, a technique for establishing a masked hearing threshold which is an intrinsically complex concept.

Another suggestion made by the participants is that there could be a ‘Help’ link, where more in-depth information could be accessed on demand. This might include a glossary of terms, explanations about why some aspects of the assessment are done (for example, why you would perform bone conduction), or a link to the corresponding part of the clinical protocol document. Hyperlinks or context-sensitive search functions may facilitate students to focus on the most relevant learning points (Huwendick et al., 2009a). It would also be helpful if there was an ‘Undo’ option, to re-set the audiometric settings to before the error was made, addressing the problem that commonly occurred of becoming stuck in a feedback loop. Alternatively there could be an ‘Ignore’ or ‘Override’ function, allowing the user the choice of either making corrections or continuing regardless of the feedback, at the risk of getting it wrong. Lastly, there needs to be an additional feedback message to remind the user to mark the threshold on the audiogram before continuing to the next test frequency; failure to do so resulted in the system detecting an error, but repeat testing to re-obtain the threshold also triggered an error message. The system would benefit from such modifications which would aim to reduce the reported frustration experienced by some of the participants in this study, hopefully increase the efficiency of use and preserve the beneficial nature of training on the CAS with formative feedback.

### 4.6.2 Further Research

Future research that would extend the findings of this study and contribute to on-going software development might include conducting a similar study with a cohort of MAud students, being the target population for the CAS. Whilst the benefits of simulation for clinical training are well established, simulation is up-and-coming in the domain of audiology, and many of the design intricacies embedded in simulators, such as feedback, are yet to be evaluated with respect to their effect on learning. It is also recognised that, as the experience of students change from novice to advanced, their support requirements change from dependant learners, needing specific guidance to independent, self-efficacious practitioners. Two types of feedback are now available in the CAS: the formative feedback training mode, and the original summative feedback condition. Studying the effect of feedback on learning for audiology students specifically, as well as the effect of feedback at different stages during their training would provide findings to guide the utilisation of the CAS throughout Audiology training in addition to recommendations for refinement of Audiology simulation software. In addition, the CAS may prove beneficial to learning for students outside of the MAud programme. Practical training with the CAS could be assimilated into undergraduate CMDS courses, such as CMDS 242: Introduction to Audiology. Future research may examine how such a system is best utilised within the syllabus and the effect it has on learning in addition to the traditional course content.

## 4.7 Clinical Implications

The findings of this study provide objective evidence that the formative feedback training version of the CAS is beneficial to learning in novice students, and relieves cognitive load in subsequent assessment. It would be a useful tool for beginning Audiology students and for students in the CMDS 242 course for gaining practical experience in with a variety of aetiologies, without placing demand on limited clinic space or risking patient safety. Whilst simulation alone is not considered an adequate substitute for real patient interaction, the opportunity for learner-oriented practice without time constraints alongside the provision of objective, immediate feedback within a self-contained and highly portable package offers an attractive supplement to traditional teaching methods. Figure 13 displays a proposed model of integrating the CAS with the different types of feedback into the undergraduate and postgraduate Audiology courses where there is an emphasis on clinical practice.

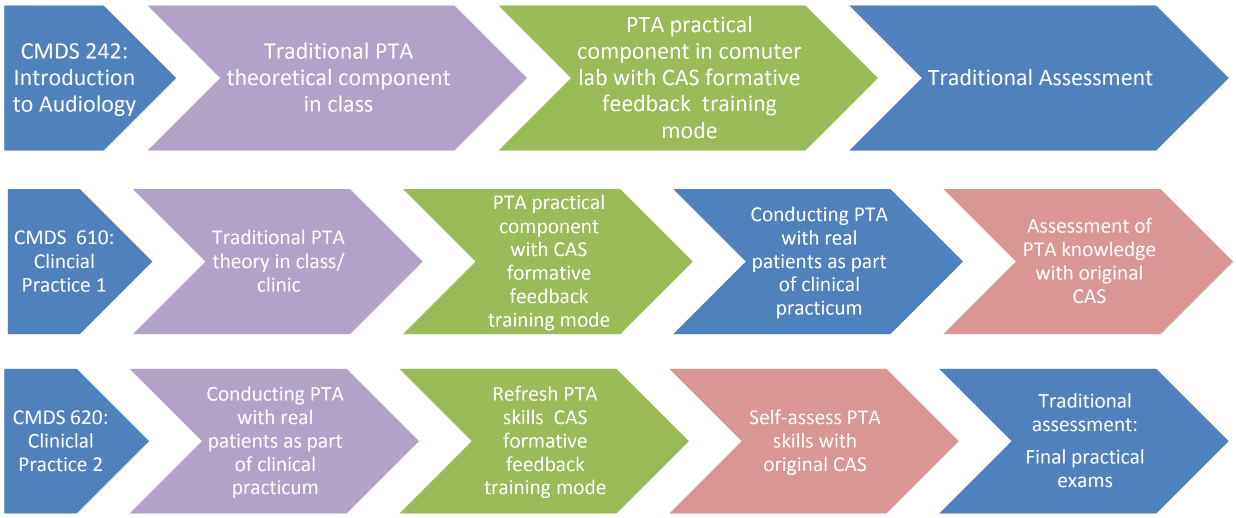



Figure 13: Model of integrating CAS into Audiology courses

## 4.8 Conclusion

The use of simulation in clinical training has become more popular with advances in technology, driven by increasingly limited clinical spaces despite growing demands for safe and extensive clinical experiences. A simulator specifically designed for clinical audiology training was developed at the University of Canterbury, which required a more supportive from of feedback embedded in the software in order to be of increased benefit for novice students. The current study measured the effects of a formative feedback training version of the CAS on learning and cognitive load in subjects with relatively little audiology experience. Formative feedback had a large positive effect on learning in comparison to PTA training on the CAS with the original summative feedback. Formative feedback was observed to increase and arguably optimise perceived workload during training, yet reduce during assessment. The opposite was seen for those who received summative feedback. Participants provided some recommendations for modification of the formative feedback training mode, but generally reacted to the CAS positively and recognised the benefits of utilising the software in both undergraduate and post-graduate Audiology courses. Future research could be directed at the integration of the CAS as a clinical training package in conjunction with traditional teaching methods, monitoring the impact on learning and enhancement of clinical abilities.

# References

American Speech-Language-Hearing Association. (2005). *Guidelines for Manual Pure-Tone Threshold Audiometry* [Guidelines]. Available from www.asha.org/policy.

American Speech-Language-Hearing Association. (2007). Scope of Practice in Speech-Language Pathology [Scope of Practice]. Available from www.asha.org/policy.

Archer, J. C. (2010). State of the science in health professional education. Effective feedback. *Medical Education, 44*(1), 101–108.

Bangert-Drowns, R., Kulik, C., Kulik. J., & Morgan, M. (1991). The Instructional Effect of Feedback in Test-like Events. *Review of Educational Research, 61*(2), 213–238.

Botezatu, M., Hult, H., Tessma, M., & Fors, U. (2010). Virtual patient simulation for learning and assessment: Superior results in comparison with regular course exams. *Medical Teacher, 32*(10), 845–850.

Bradley, P. (2006). The history of simulation in medical education and possible future directions. *Medical Education, 40*(3), 254–262.

Bransford, J., Brown, A., & Cocking, R. (Eds.). (2000). How people learn: Brain, mind, experience, and school: Expanded edition. Washington, DC: The National Academies Press.

Butler, A., Godbole, N., & Marsh, E. (2013). Explanation feedback is better than correct answer feedback for promoting transfer of learning. *Journal of Educational Psychology (in press).* Retrieved from: http://people.duke.edu/~ab259/pubs/ButlerGodbole&Marsh%282013%29.pdf

Cantillon, P., & Sargeant, J. (2008). Giving feedback in clinical settings. *British Medical Journal*, 337:a1961.

Carhart, R. and Jerger J. (1959). Preferred Methods for Clinical Determination of Pure-Tone Thresholds. Journal of Speech, Language, and Hearing Research, 24, 330-

345.

Chan, J., & Lam, S. (2010). Effects of different evaluative feedback on students' self-efficacy in learning. *Instructional Science, 38*, 37–58.

Chase, J. R., & Houmanfar, R. (2009). The differential effects of elaborate feedback and basic feedback on student performance in a modified, personalized system of instruction course. *Journal of Behaviour Education, 18,* 245–265.

Cohen, J. (1988). Statistical Power Analysis for the Behavioural Sciences. (2nd ed). Hillsdale, NJ: Lawrence Erlbaum Associates

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science, 1*(3), 98–101.

Cook, D., & Triola, M. (2009). Virtual patients: A critical literature review and proposed next steps. *Medical Education, 43*(4), 303–311.

Cotton, K. (1988). Monitoring student learning in the classroom. School improvement research series: Research you can use. Office of Educational Research and Improvement (OERI), U.S. Department of Education. Retrieved from: http://educationnorthwest.org/webfm_send/541

Cross, S. E., & Madson, L. (1997). Models of the self: self-construals and gender. *Psychological Bulletin*, *122*(1), 5–37.

Dempsey, J., & Sales, G. (Eds.). (1993). Interactive Instruction and Feedback. Englewood Cliffs, NJ: Educational Technology Publications Inc.

Djamasbi, S., & Loiacono, E. (2008). Do men and women use feedback provided by their Decision Support Systems (DSS) differently? *Decision Support Systems*, *44*(4), 854–869.

Dünser, A., Heitz, A., & Moran, C. (2010). Development of a virtual Avatar-Based application for audiology training. In *15th Annual Cyber Therapy & Cyber*

*Psychology Conference*.

Duvivier, R., Dalen, J. van, Muijtjens, A., Moulaert, V., Van der Vleuten, C., Scherpbier, A. (2011). The role of deliberate practice in the acquisition of clinical skills. *BMC Medical Education, 11*: 101

Ellaway, R., & Masters, K. (2008). AMEE Guide 32. E-Learning in medical education. Part 1. Learning, teaching and assessment. *Medical Teacher, 30*(5), 455–473.

Ericsson, K., Krampe, R., & Tesch-Romer, C. (1993). The Role of Deliberate Practice in the Acquisition of Expert Performance. *Psychological Review, 100*(3), 363-406

Fiorella, L., Vogel-Walcutt, J., & Schatz, S. (2011). Applying the modality principle to real-time feedback and the acquisition of higher-order cognitive skills. *Educational Technology Research and Development*, *60*(2), 223–238.

Galy, E., Cariou, M., & Mélan, C. (2012). What is the relationship between mental workload factors and cognitive load types? *International Journal of Psychophysiology: official journal of the International Organization of Psychophysiology*, *83*(3), 269–75.

Good, M. L. (2003). Patient simulation for training basic and advanced clinical skills. *Medical Education, 37*(Suppl. 1), 14–21.

Hart, S., & Staveland, L. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Retrieved from: http://humanfactors.arc.nasa.gov/groups/TLX/downloads/NASA-TLXChapter.pdf

Hauer, K., & Kogan, J. (2012). Realising the potential value of feedback. *Medical education*, *46*(2), 140–2.

Heitz, A. (submission pending, March 2013). Assessing Virtual Patient-based Simulations for Clinical Education (Unpublished doctoral thesis). University of Canterbury, Christchurch, New Zealand.

Hoffmann, G., & Tonshoff, B. (2009). Design principles for virtual patients. A focus group study among students. *Medical Education, 43*, 580–588.

Howland, S. (2012). Immersive Education: Virtual reality in Clinical Audiology: A pilot study of the effectiveness of a new patient simulator program on audiology students' performance on case history tasks (Unpublished master's thesis). University of Canterbury, Christchurch, New Zealand.

Huwendiek, S., Reichert, F., Bosse, H-M., de Leng, B., van der Vleuten, C., Haag, M., … Tonshoff, B. (2009a). Design principles for virtual patients: a focus group study among students. *Medical Education, 43*, 580–588.

Huwendiek, S., Leng, B., Zary, N., Fischer, M., Ruiz, J., & Ellaway, R. (2009b). Towards a typology of virtual patients. *Medical Teacher, 31*, 743–748.

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, *77*(1), 81–112.

Huxham, M. (2007). Fast and effective feedback: are model answers the answer? *Assessment & Evaluation in Higher Education*, *32*(6), 601–611.

Issenberg, S., McGaghie, W., Petrusa, E., Gordon, L., & Scalese, R. (2005). Features and uses of high fidelity medical simulations that lead to effective learning. A BEME systematic review. *Medical Teacher, 27*(1), 10–28.

Jong, T. de (2009). Cognitive load theory, educational research, and instructional design: some food for thought. *Instructional Science*, *38*(2), 105–134.

Keeley, P., Eberle, F., & Farrin, L. (2005). Formative assessment probes: uncovering student's ideas in science. *Science Scope, 28*(4). 18–21.

Kluger, A. N., & Denisi, A. (1996). The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory. *Psychological Bulletin*, *II*(2), 254–284.

Kulhavy, R., & Stock, W. (1989). Feedback in Written Instruction: The Place of

Response Certitude. *Educational Psychology Review, 1*(4), 279–308.

Kulik, J., & Kulik, C. (1988). Timing of feedback and verbal learning. *Review of Educational Research, 58*(1), 79–97.

Mason, B. and Bruning, R. (2001) *Providing Feedback in Computer-based Instruction: What the Research Tells Us.* Retrieved from: http://dwb.unl.edu/Edit/MB/MasonBruning.html.

Mayer, R. (2008). *Learning and Instruction.* Upper Saddle River, NJ: Pearson Education, Inc.

Merrienboer, J. van, & Sweller, J. (2005). Cognitive Load Theory and Complex Learning: Recent Developments and Future Directions. *Educational Psychology Review, 17*, 147–177.

Miller, A. S. (2001). Literature review: Workload Measures. National Advanced Driving Simulator, University of Iowa, USA. Retrieved from: http://www.nads-sc.uiowa.edu/publicationStorage/200501251347060.N01-006.pdf .

Moreno, R., & Mayer, R. (2007). Interactive Multimodal Learning Environments. *Educational Psychology Review*, *19*(3), 309–326.

Moreno, R., Mayer, R., & Lester, J. (2001). The Case for Social Agency in Computer-Based Teaching : Do Students Learn More Deeply When They Interact With Animated Pedagogical Agents. *Cognition and Instruction, 19*(2), 177–213.

Moreno, R. (2004). Decreasing cognitive load for novice students: Effects of explanatory versus corrective feedback in discovery-based multimedia. *Instructional Science, 32*, 99–113.

Mumm, J., & Mutlu, B. (2011). Designing motivational agents: The role of praise, social comparison, and embodiment in computer feedback. *Computers in Human Behavior*, *27*(5), 1643–1650.

Human Performance Research Group. (1998). NASA Task Load Index (TLX): Paper and

Pencil Package. Moffett Field, CA: NASA-Ames Research Center. Retrieved from: http://humansystems.arc.nasa.gov/groups/TLX/downloads/TLX_pappen_manual.pdf

New Zealand Audiological Society. (2012). Best Practice Guidelines: Adult Pure-tone Audiometry. Retrieved from: http://www.audiology.org.nz/mnzas-standards-of-practice.aspx

Osman, M. (2010). Virtual Tutoring. An online environment for scaffolding students' metacognitive problem solving expertise. *Journal of Turkish Science Education, 7*(4), 3–12.

Paas, F., & Merrienboer, J. van (1994). Variability of worked examples and transfer of geometrical problem solving skills: A cognitive load approach. *Journal of Educational Psychology, 86,* 122–133.

Pallant, J. (2005). SPSS Survival Manual: A step by step guide to data analysis using SPSS for Windows (Version 12). Allen & Unwin; New South Wales.

Plass, J., Homer, B., & Hayward, E. (2009). Design factors for educationally effective animations and simulations. *Journal of Computing in Higher Education*, *21*(1), 31–61.

Rosen, K. (2008). The history of medical simulation. *Journal of Critical Care, 23*(2), 157–166.

Schroth, M. (1992). The effects of delay of feedback on a delayed concept formation transfer task. *Contemporary Educational Psychology, 17*(1), 78–82.

Shute, V. (2008). Focus on Formative Feedback. *Review of Educational Research,* 78(1): 153–189.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257–285.

Sweller, J., Merrienboer, J. van, & Paas, F. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review, 10*(3), 251–296.

University of Canterbury Speech & Hearing Clinic. (2012). Audiology protocols and guidelines. *Unpublished*.

van de Ridder, J., Stokking, K., McGaghie, W., & ten Cate, O. (2008). What is feedback in clinical education? *Medical education*, *42*(2), 189–97.

White, K. & Weight, B. (Eds). (2000). *The online teaching guide: A handbook of attitudes, strategies and techniques for the virtual classroom*. Boston, MA: Allyn and Bacon.

Wolf, M. (1978). Social validity. The case for subjective measurement or how applied behaviour analysis is finding its heart. *Journal of Applied Behavior Analysis, 11*(2) 203–214.

Ziv, A., Ben-David, S., & Ziv, M. (2005). Simulation based medical education. An opportunity to learn from errors. *Medical Teacher 27*(3), 193–199.

# Appendices

## Appendix I

Formative feedback guidelines from Shute (2008)

*Formative Feedback Guidelines to Enhance Learning (Things to Do)*

| | Prescription | Description and references |
|---|---|---|
| 1 | Focus feedback on the task, not the learner. | Feedback to the learner should address specific features of the learner's work in relation to the task, with suggestions on how to improve (e.g., Butler, 1987; Corbett & Anderson, 2001; Kluger & DeNisi, 1996; Narciss & Huth, 2004). |
| 2 | Provide elaborated feedback to enhance learning. | Feedback should describe the what, how, and/or why of a given problem. This type of cognitive feedback is typically more effective than verification of results (e.g., Bangert-Drowns et al., 1991; Gilman, 1969; Mason & Bruning, 2001; Narciss & Huth, 2004; Shute, 2006). |
| 3 | Present elaborated feedback in manageable units. | Provide elaborated feedback in small enough pieces so that it is not overwhelming and/or discarded (Bransford et al., 2000; Sweller et al., 1998). Presenting too much information may not only result in a superficial learning, but also invoke cognitive overload (e.g., Mayer & Moreno, 2002; Phye & Bender, 1989). A stepwise presentation of feedback offers the possibility to control for mistakes and gives learners sufficient information to correct errors on their own. |
| 4 | Be specific and clear with feedback messages. | If feedback is not specific or clear, it can impede learning and can frustrate learners (e.g., Moreno, 2004; Williams, 1997). If possible, try to link feedback clearly and specifically to goals and performance (Hoska, 1993; Song & Keller, 2001). |
| 5 | Keep feedback as simple as possible but no simpler (based on learner needs and instructional constraints). | Simple feedback is generally based on one cue (e.g., verification or hint) and complex feedback on multiple cues (e.g., verification, correct response, error analysis). Keep feedback as simple and focused as possible. Generate only enough information to help students and not more. Kulhavy et al. (1985) found that feedback that was too complex did not promote learning compared to simpler feedback. |
| 6 | Reduce uncertainty between performance and goals. | Formative feedback should clarify goals and seek to reduce or remove uncertainty in relation to how well learners are performing on a task and what needs to be accomplished to attain the goal(s) (e.g., Ashford, Blatt, & VandeWalle, 2003; Bangert-Drowns et al., 1991). |
| 7 | Give unbiased, objective feedback, written or via computer. | Feedback from a trustworthy source will be considered more seriously than other feedback, which may be disregarded. This may explain why computer-based feedback is often better than human-delivered in some experiments in that perceived biases are eliminated (see Kluger & DeNisi, 1996). |
| 8 | Promote a learning goal orientation via feedback. | Formative feedback can be used to alter goal orientation—from a focus on performance to a focus on learning (Hoska, 1993). This can be facilitated by crafting feedback emphasizing that effort yields increased learning and performance and that mistakes are an important part of the learning process (Dweck, 1986). |
| 9 | Provide feedback after learners have attempted a solution. | Do not let learners see answers before trying to solve a problem on their own (i.e., presearch availability). Several studies that have controlled presearch availability show a benefit of feedback, while studies without such control show inconsistent results (Bangert-Drowns et al., 1991). |

*Formative Feedback Guidelines to Enhance Learning (Things to Avoid)*

| | Prescription | Description and references |
|---|---|---|
| 10 | Do not give normative comparisons. | Feedback should avoid comparisons with other students—directly or indirectly (e.g., grading on the curve). In general, do not draw attention to self during the course of learning (Kluger & DeNisi, 1996; Wiliam, in press). |
| 11 | Be cautious about providing overall grades. | Feedback should note areas of strength and provide information on how to improve, as warranted and without overall grading. Wiliam (in press) summarized the following findings: (a) students receiving just grades showed no learning gains, (b) those getting just comments showed large gains, and (c) those with grades and comments showed no gains (likely due to focusing on the grade and ignoring comments). Effective feedback relates to the content of the comments (Butler, 1987; McColskey & Leary, 1985). |
| 12 | Do not present feedback that discourages the learner or threatens the learner's self-esteem. | This prescription is based not only on common sense, but also research reported in Kluger and DeNisi (1996), which reports feedback interventions that undermine learning as it draws focus to the self and away from the task at hand. In addition, do not provide feedback that is either too controlling or critical of the learner (Baron, 1993; Fedor, Davis, Maslyn, & Mathieson, 2001). |
| 13 | Use praise sparingly, if at all. | Kluger & DeNisi (1996), Butler (1987), and others have noted that use of praise as feedback directs the learner's attention to self, which distracts from the task and consequently from learning. |
| 14 | Try to avoid delivering feedback orally. | This also was addressed in Kluger & DeNisi (1996). When feedback is delivered in a more neutral manner (e.g., written or computer-delivered), it is construed as less biased. |
| 15 | Do not interrupt the learner with feedback if the learner is actively engaged. | Interrupting a student who is immersed in a task—trying to solve a problem or task—can be disruptive to the student and impede learning (Corno & Snow, 1986). |
| 16 | Avoid using progressive hints that always terminate with the correct answer. | While hints can be facilitative, they can also be abused. If they are employed to scaffold learners, make provisions to prevent their abuse (e.g., Aleven & Koedinger, 2000; Shute, Woltz, & Regian, 1989). Consider using prompts and cues (i.e., more specific kinds of hints). |
| 17 | Do not limit the mode of feedback presentation to text. | Exploit the potential of multimedia to avoid cognitive overload due to modality effects (e.g., Mayer & Moreno, 2002). Do not default to presenting feedback messages as text. Instead, consider alternative modes of presentation (e.g., acoustic, visual). |
| 18 | Minimize use of extensive error analyses and diagnosis. | The cost of conducting extensive error analyses and cognitive diagnosis may not provide sufficient benefit to learning (Sleeman et al., 1989; VanLehn et al., 2005). Furthermore, error analyses are rarely complete and not always accurate, thus only are helpful in a subset of circumstances. |

*Formative Feedback Guidelines in Relation to Timing Issues*

| | Prescription | Description and references |
|---|---|---|
| 19 | Design timing of feedback to align with desired outcome. | Feedback can be delivered (or obtained) either immediately after some activity or delayed. Immediate feedback can help fix errors in real-time, producing greater immediate gains and more efficient learning, (Corbett & Anderson, 2001; Mason & Bruning, 2001), but delayed feedback has been associated with better transfer of learning (e.g., Schroth, 1992). |
| 20 | For difficult tasks, use immediate feedback. | When a student is learning a difficult new task (where difficult is relative to the learner's capabilities), it is better to use immediate feedback, at least initially (Clariana, 1990). This provides a helpful safety net so the learner does not get bogged down and/or frustrated (Knoblauch & Brannon, 1981). |
| 21 | For relatively simple tasks, use delayed feedback. | When a student is learning a relatively simple task (again, relative to capabilities), it is better to delay feedback to prevent feelings of feedback intrusion and possibly annoyance (Clariana, 1990; Corno & Snow, 1986). |
| 22 | For retention of procedural or conceptual knowledge, use immediate feedback. | In general, there is wide support for use of immediate feedback to promote learning and performance on verbal, procedural, and even tasks requiring motor skills (Anderson et al., 2001; Azevedo & Bernard, 1995; Corbett & Anderson, 1989, 2001; Dihoff et al., 2003; Phye & Andre, 1989). |
| 23 | To promote transfer of learning, consider using delayed feedback. | According to some researchers (e.g., Kulhavy et al., 1985; Schroth, 1992), delayed may be better than immediate feedback for transfer task performance, although initial learning time may be depressed. This needs more research. |

*Formative Feedback Guidelines in Relation to Learner Characteristics*

| | Prescription | Description and references |
|---|---|---|
| 24 | For high-achieving learners, consider using delayed feedback. | Similar to the Clariana (1990) findings cited in Table 4, high-achieving students may construe a moderate or difficult task as relatively easy and hence benefit by delayed feedback (see also Gaynor, 1981; Roper, 1977). |
| 25 | For low-achieving learners, use immediate feedback. | The argument for low-achieving students is similar to the one above, only these students need the support of immediate feedback in learning new tasks they may find difficult (see Gaynor, 1981; Mason & Bruning, 2001; Roper, 1977). |
| 26 | For low-achieving learners, use directive (or corrective) feedback. | Novices, or struggling students, need support and explicit guidance during the learning process (Knoblauch & Brannon, 1981; Moreno, 2004), thus hints may not be as helpful as more explicit, directive feedback. |
| 27 | For high-achieving learners, use facilitative feedback. | Similar to the above, high-achieving students or more motivated ones benefit from feedback that challenges them, such as hints, cues, and prompts (Vygotsky, 1987). |
| 28 | For low-achieving learners, use scaffolding. | Provide early support and structure for low-achieving students (or those with low self-efficacy) to improve learning and performance (e.g., Collins et al., 1989; Graesser, McNamara, & VanLehn, 2005). |
| 29 | For high-achieving learners, verification feedback may be sufficient. | Hanna (1976) presented findings that suggest that high-achieving students learn more efficiently if permitted to proceed at their own pace. Verification feedback provides the level of information most helpful in this endeavor. |
| 30 | For low-achieving learners, use correct response and some kind of elaboration feedback. | Using the same rationale as with supplying scaffolding to low-achieving students, the prescription here is to ensure low-achieving students receive a concrete, directive form of feedback support (e.g., Clariana, 1990; Hanna, 1976). |
| 31 | For learners with low learning orientation (or high performance orientation), give specific feedback. | As described in the study by Davis et al., (2005), if students are oriented more toward performance (trying to please others) and less toward learning (trying to achieve an academic goal), provide feedback that is specific and goal directed. Also, keep the learner's eye on the learning goal (Hoska, 1993). |

## Appendix II

Case information, audiograms and notes for software developers

# AUDIOGRAMS/INFORMATION FOR CASES

## Case 1

**Focus**: Transducer selection; threshold seeking procedure; hearing testing with normal hearing

**Patient**: age 30/Female

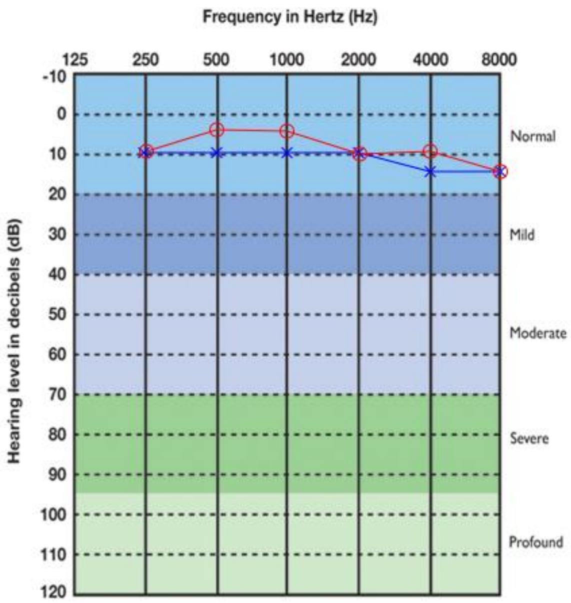
**Case background**: GP prescribed drops for Otitis Externa (outer ear infection); ears feel itchy and were painful but feeling better with the drops. Concerned her hearing is being affected.

** **Otoscopy**: Red, irritated/weeping ear canals bilaterally Alex/Jakob: can we have this sentence inserted as a comment to accompany the otoscopy picture?

**Test**: PTA

** **Transducer**: Headphones
Inserts selected = **Error**
**TRANSDUCER SELECTION RULE:** Use the information obtained from otoscopy to judge what type of transducer to use. Default to insert earphones, unless there is an outer ear infection, grommets or perforated ear drum, where supra aural head phones are used
Bone conductor selected = **Error**
**AC BEFORE BC RULE:** Air conduction thresholds are obtained first, where the stimulus or tone is presented via supra-aural headphones or insert earphones. If a hearing loss is evident (i.e. thresholds that are greater than 20dB HL) the bone conductor is used to differentiate the type of hearing loss (i.e. conductive or sensorineural).

**Reliability**: V. Good




Procedure for obtaining thresholds - Apply the following rules (see case 2 and rules summary document):

- AC EAR SELECTION RULE
- TEST FREQUENCY ORDER RULE

- INITIAL PRESENTATION LEVEL RULE
- 10dB DOWN RULE
- 5dB UP RULE
- 2 ASCENDING RUNS RULE
- MARK THRESHOLD RULE
- COMPLETE EAR INFO RULE
- SUBSEQUENT PRESENTATION LEVEL RULE

Examples of Right ear start levels and threshold seeking procedures:

| Dial level (dB HL) | Response | Action | Dial level (dB HL) | Response | Action |
|---|---|---|---|---|---|
| 1000Hz | | | 2000Hz | | |
| 30 | Y | down 10 | 20 | Y | down 10 |
| 20 | Y | down 10 | 10 | Y | down 10 |
| 10 | Y | down 10 | 0 | N | up 5 |
| 0 | N | up 5 | 5 | N | up 5 |
| 5 | Y | down 10 | 10 | Y | down 10 |
| -5 | N | up 5 | 0 | N | up 5 |
| 0 | N | up 5 | 5 | N | up 5 |
| 5 | Y | mark t/h | 10 | Y | mark t/h |
| 4000Hz | | | 8000Hz | | |
| 30 | Y | down 10 | 30 | Y | down 10 |
| 20 | Y | down10 | 20 | Y | down 10 |
| 10 | Y | down 10 | 10 | N | up 5 |
| 0 | N | up 5 | 15 | Y | down 10 |
| 5 | N | up 5 | 5 | N | up 5 |
| 10 | Y | down 10 | 10 | N | up 5 |
| 0 | N | up 5 | 15 | Y | mark t/h |
| 5 | N | up 5 | | | |
| 10 | Y | mark t/h | | | |
| 500Hz | | | 250 Hz | | |
| 25 | Y | down 10 | 20 | Y | down 10 |
| 15 | Y | down 10 | 10 | Y | down 10 |
| 5 | Y | down 10 | 0 | N | up 5 |
| -5 | N | up 5 | 5 | N | up 5 |
| 0 | N | up 5 | 10 | Y | down 10 |
| 5 | Y | down 10 | 0 | N | up 5 |
| -5 | N | up 5 | 5 | N | up 5 |
| 0 | N | up 5 | 10 | Y | mark t/h |
| 5 | Y | mark t/h | | | |

- Repeat for L) ear (Let me know if you need examples)

## Case 2

**Focus**: Threshold seeking procedure; tone presentation (2 second duration and non-predictive); testing inter-octaves

**Patient**: age 48/Male

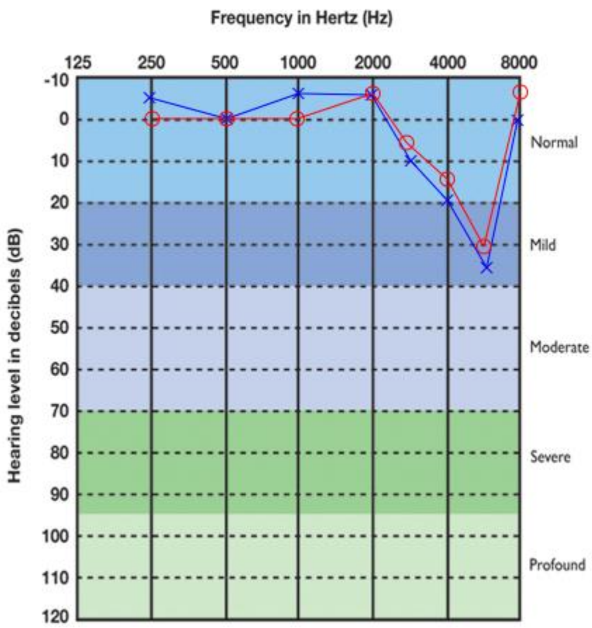
**Case background**: Recently noticed it is becoming harder to hear in background noise and on the phone; has own business as an arbourist (lots of chainsawing and industrial mulching), and is a keen hunter (deer).

**Otoscopy**: Clear ear canals and ear drums bilaterally

**Test**: PTA

**Transducer**: Inserts (apply TRANSDUCER SELECTION RULE)

**Reliability**: V. Good




** Procedure instructions:

- start with R ear

Left ear selected = **Error - EAR SELECTION RULE**: start testing with the reported better ear; if both are the same, start with the right ear

- start at 1000Hz

Any other frequency = **Error - TEST FREQUENCY ORDER RULE**: start at 1000Hz, ascend up to 2000, 4000 and 8000Hz, then descend to 500 and 250Hz

- present a 30dB tone

<30dB = **Error – INITIAL PRESENTATION LEVEL RULE**: the level of the very first tone presented should be well above the patient's actual threshold so they become familiar with what they need to listen for. As a general rule, start at 30dB. If the patient doesn't respond to the first tone, increase in 15-20dB steps until they respond

- the following threshold seeking procedure should be followed, with the corresponding responses:

30dB HL – Y → drop 10
- anything other than a 10dB drop = **Error – 10dB DOWN RULE**: when the patient responds to a tone, decrease the signal level by 10dB, and present the tone again. Repeat descending in 10dB steps and presenting the tone until the patient doesn't respond.

20dB HL – Y → drop 10

10dB HL – Y → drop 10

0dB HL – Y → drop 10

-10dB HL – N → up 5dB
- repeat at -10/up by anything other than 5dB = **Error – 5dB UP RULE**: when the patient doesn't respond to a tone, increase by 5dB and present the tone again. Repeat ascending in 5dB steps and presenting the tone until the patient does respond

-5dB HL – N → up 5dB

0dB HL - Y = response 1 on ascending run → drop 10
- mark threshold = **Error – 2 ASCENDING RUNS RULE**: the threshold is the level where a response is obtained for <u>two ascending runs</u> using the 10dB down, 5dB up method.
Also, apply 10dB DOWN RULE, if level change does not equal a change of -10dB.

-10dB HL – N → up 5

-5dB HL – N → up 5

0dB HL – Y = response 2 on ascending run → Mark threshold at 5dB for 1000Hz.
- anything other than 'mark threshold' = **Error - MARK THRESHOLD RULE**: two ascending runs have been obtained. Mark the threshold at the level where the response was obtained for the two ascending runs.

- stay on R ear
- switched to Left ear = **Error - COMPLETE AC EAR INFO RULE**: obtain all air conduction thresholds for this ear before switching to the other ear

- ascend to 2000Hz (apply TEST FREQUENCY ORDER RULE)

- present a 20dB tone
>/< previous threshold + 15-20dB = **Error - SUBSEQUENT PRESENTATION LEVEL RULE:** once the first threshold is determined, the start presentation level at the next frequency can be 15-20dB higher than the previous threshold

- the following tone intensities should be presented, with the following responses:

| Dial level (dB HL) | Response | Action |
|---|---|---|
| 20 | Y | down 10 |
| 10 | Y | down 10 |
| 0 | Y | down 10 |
| -10 | N | up 5 |
| -5 | N | up 5 |
| 0 | Y | down 10 |
| -10 | N | up 5 |
| -5 | N | up 5 |
| 0 | Y | mark t/h |

- repeat for 4000, 8000, 500 and 250Hz (do you need these tabulated as well?)

- thresholds for 1500, 3000 and 6000Hz should also be obtained
- Switches to left ear without testing1500, 3000 and 6000Hz = **Error:**
**TEST INTEROCTAVES RULE**: when noise induced hearing loss is suspected, hearing thresholds for inter-octave frequencies of 1500, 3000 and 6000Hz should be obtained. Also test inter-octaves (750, 1500, 3000, 6000Hz) if a more detailed audiogram is required, such as when a patient complains of hearing difficulties but the audiogram seems normal. The inter-octave should also be tested if the AC threshold drops 20dB between octave frequencies.

- Once a full audiogram is obtained for the right ear, switch to the left.
All correct upon switching ears → **positive f/b**, e.g. smiley face/happy ear/ "Good job, this ear is correct"…?

- Start at 1000Hz, at a level of 30dB; repeat threshold seeking procedure as above; continue obtaining thresholds for the test frequencies including interoctaves as per TEST INTEROCTAVES RULE

Other possible errors:
- BC selected to test 500-4000Hz; BC selected to test <500Hz or >4000Hz = **Error**:
**WHEN TO DO BC RULE**: Obtain bone conduction thresholds at octave frequencies from 500 to 4000Hz when air conduction thresholds at these frequencies are outside of normal limits (i.e. =/>20dB HL) to differentiate the type of hearing loss (i.e. conductive or sensorineural). Only 500 – 4000Hz can be tested reliably with bone conduction

## Case 3

**Focus**: Masking – BC/Air-bone gap

**Patient**: age 51/Female

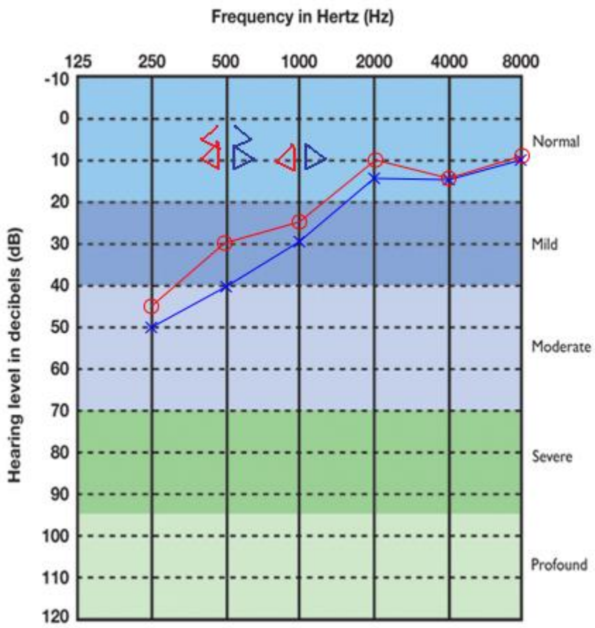
**Case background**: Noticed her hearing 1st started to reduce when she had her 2 children in her 30s, has got worse over time; feels she has no trouble provided everything is louder – phone, TV, etc. Her mother was always hard of hearing, and her grandmother also – thought to have been 'deaf' from their 20s.

**Otoscopy**: Clear ear canals and ear drums bilaterally

**Test**: PTA

**Transducer**: Inserts (apply TRANSDUCER SELECTION RULE & AC BEFORE BC RULE)

**Reliability**: V. Good




Procedure:
- start with R ear (apply AC EAR SELECTION RULE)

- start frequency: 1000Hz (apply TEST FREQUENCY ORDER RULE)

- start presentation level: 30dB (apply INITIAL PRESENTATION LEVEL RULE)

- seek threshold at this frequency (apply the following rules: 10dB DOWN RULE 5dB UP RULE, 2 ASCENDING RUNS RULE, MARK THRESHOLD RULE)

- repeat for 2000, 4000, 8000, 500 & 250Hz (apply SUBSEQUENT PRESENTATION LEVEL RULE; TEST FREQUENCY ORDER RULE)

- Switch ears to obtain AC thresholds for the L) ear (as above)

** BC thresholds need to be obtained for 500Hz & 1kHz for both ears (apply WHEN TO DO

Occlusion Effect

BC RULE)

- **BC EAR SELECTION RULE**: For bone conduction, start with the *worst* hearing ear first.
→ L) ear then R) ear

- **START LEVELS FOR BC RULE**: The start level for BC audiometry should be 10-15 dB above the air conduction threshold of the *better* ear
→ BC start level L) ear: 1000Hz = 35dB HL; 500Hz = 40dB HL; same for R) ear

- BC thresholds are down 10/up 5 threshold found with the usual seeking procedure (apply the following: 10dB DOWN RULE, 5dB UP RULE, 2 ASCENDING RUNS RULE, MARK THRESHOLD RULE)
→ e.g. for the L) ear at 1000Hz:

| Dial level (dB HL) | Response | Action |
|---|---|---|
| 35 | Y | down 10 |
| 25 | Y | down 10 |
| 15 | Y | down 10 |
| 5 | N | up 5 |
| 10 | Y | down 10 |
| 0 | N | up 5 |
| 5 | N | up 5 |
| 10 | Y | mark t/h |

! Let me know if you need more e.gs

- BC thresholds at 500 & 1000Hz need to be masked:
**WHEN TO MASK BC RULE**: A masked bone conduction response is required when the air conduction threshold in the test ear is greater or more severe than the unmasked bone conduction threshold (at the same frequency in the same ear), by =/> 15dB (i.e. a significant air-bone gap).

- masking noise is played into the non-test ear:
**WHAT EAR TO MASK RULE:** Narrow band masking noise is used to mask pure tones. This is played into the non test ear via the air conduction pathway whilst true masked thresholds are obtained from the test ear.

- masking noise is presented at the calculated initial masking level:
**INITIAL MASKING LEVEL RULE**: The level of noise initially played into the non test ear at a given frequency is calculated with the following: level of the threshold in the non test ear + 10dB safety factor for guaranteed effective masking + occlusion effect at low frequencies

| frequency | 250Hz | 500Hz | 1000Hz | 2000Hz | 4000Hz | 8000z |
|---|---|---|---|---|---|---|
| supra-aural headphones | 30dB | 20dB | 10dB | - | - | - |
| inserts | 10dB | 10dB | - | - | - | - |

IML for R) ear when getting masked BC t/h in L) ear:
1000Hz: 25 dB HL t/h + 10dB safety + 10dB occlusion effect = 45dB EM (effective masking)
500Hz: 30 dB HL t/h + 10dB safety + 20dB occlusion effect = 60dB EM

IML for L) ear when getting masked BC t/h in R) ear:
1000Hz: 30 dB HL t/h + 10dB safety + 10dB occlusion effect = 50dB EM
500Hz: 40 dB HL t/h + 10dB safety + 20dB occlusion effect = 70dB EM

- true bone conduction thresholds are sought from the test ear by using the plateau method, and marked on the audiogram accordingly:

**PLATEAU RULE**: Whilst masking noise is played into the non test ear, present the tone to the test ear at the level of the unmasked threshold. Increase the level of masking noise in 10dB steps, and present the tone in the test ear until the masking has been increased by 20dB (2 steps of 10 dB) and the patient continues to respond to the tone. Increase test tone in steps of 5dB if required for it to be audible to the patient above the masking noise. The patient's masked threshold is the level where they respond to the test tone when masking noise is increased by 20dB (i.e. a 20dB plateau)

| Masking level (ML) dB EM | Test tone (TT) dB HL | Response | Action | Masking level (ML) dB EM | Test tone (TT) dB HL | Response | Action |
|---|---|---|---|---|---|---|---|
| R) | L) | 1000Hz | | R) | L) | 500Hz | |
| 45 | 10 | Y | ML up 10 | 60 | 5 | N | TT up 5 |
| 55 | 10 | Y | ML up 10 | 60 | 10 | Y | ML up 10 |
| 65 | 10 | Y | mark t/h | 70 | 10 | Y | ML up 10 |
| | | | | 80 | 10 | Y | mark t/h |
| L) | R) | 1000Hz | | L) | R) | 500Hz | |
| 30 | 10 | Y | ML up 10 | 40 | 5 | N | TT up 5 |
| 40 | 10 | Y | ML up 10 | 40 | 10 | Y | ML up 10 |
| 50 | 10 | Y | mark t/h | 50 | 10 | Y | ML up 10 |
| | | | | 60 | 10 | Y | mark t/h |

## Case 4

**Focus**: Masking – advanced; both AC and BC masking required**

**Patient**: age 36/Male

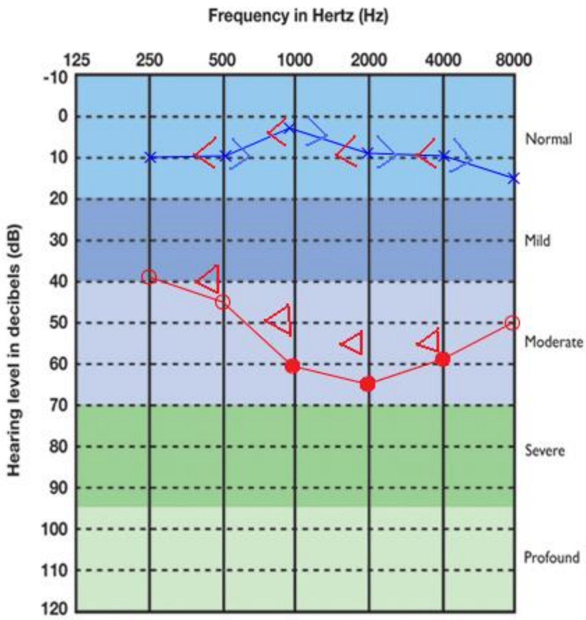
**Case background**: Woke up 2 days ago and noticed reduced hearing in his **right** ear.

**Otoscopy**: wax occlusion bilaterally

**Test**: PTA

**Transducer**: Headphones (apply TRANSDUCER SELECTION RULE & AC BEFORE BC RULE)

**Reliability**: V. Good




Procedure:

- start with L) ear (apply AC EAR SELECTION RULE)

- start frequency: 1000Hz (apply TEST FREQUENCY ORDER RULE)

- start presentation level: 30dB (apply INITIAL PRESENTATION LEVEL RULE)

- seek threshold at this frequency (apply the following rules: 10dB DOWN RULE 5dB UP RULE, 2 ASCENDING RUNS RULE, MARK THRESHOLD RULE)

- repeat for 2000, 4000, 8000, 500 & 250Hz (apply SUBSEQUENT PRESENTATION LEVEL RULE; TEST FREQUENCY ORDER RULE)

- Switch ears to obtain AC thresholds for the R) ear (as above).

Note start presentation level for the R) ear at 1000Hz, and threshold seeking procedure:

| Dial level (dB HL) | Response | Action |
|---|---|---|
| 30 | N | up 15-20 |
| 50 | N | up 15-20 |
| 70 | Y | down 10 |
| 60 | Y | down 10 |
| 50 | N | up 5 |
| 55 | N | up 5 |
| 60 | Y | down 10 |
| 50 | N | up 5 |
| 55 | N | up 5 |
| 60 | Y | mark t/h |

****** Before moving on to BC, Air conduction thresholds in the left ear need to be masked at 1000, 2000 and 4000Hz:

**WHEN TO MASK AC RULE:** Mask the air conduction threshold for the test ear if, at a given frequency, the threshold is =/> the best threshold of the non test ear at that frequency (air conduction or bone conduction threshold) **PLUS** the inter-aural attenuation value for the transducers delivering the air conduction signal (i.e. if there is a chance the non test ear could hear the level presented to the test ear).

| Inter-aural attenuation (IAA) | | | | | | |
|---|---|---|---|---|---|---|
| frequency | 250Hz | 500Hz | 1000Hz | 2000Hz | 4000Hz | 8000z |
| supra-aural headphones | 40dB | 40dB | 40dB | 40dB | 40dB | 40dB |
| inserts | 75dB | 75dB | 75dB | 50dB | 50dB | 50dB |
| bone conductor | 0dB | 0dB | 0dB | 0dB | 0dB | 0dB |

For this case:
at 1000Hz: L t/h 60dB **>** 45 (R t/h 5dB HL + 40dB IAA) → MASK
at 2000Hz: L t/h 65dB HL **>** 50 (R t/h 10dB HL + 40dB IAA) → MASK
at 4000Hz: L t/h 60dB HL **>** 50 (R t/h 10dB HL + 40dB IAA) → MASK

Apply: **INITIAL MASKING LEVEL RULE**

IML to the R) ear when obtaining masked AC t/h in the L) ear:
1000Hz: 5dB HL t/h + 10dB safety + 10dB occlusion effect = 25dB EM
2000Hz: 10dB HL + 10dB safety = 20dB EM
40000Hz: 10dB HL + 10dB safety = 20dB EM

Apply: **PLATEAU RULE**

| Masking level (ML) dB EM | Test tone (TT) dB HL | Response | Action | Masking level (ML) dB EM | Test tone (TT) dB HL | Response | Action |
|---|---|---|---|---|---|---|---|
| R) | L) | 1000Hz | | R) | L) | 2000Hz | |
| 25 | 60 | Y | ML up 10 | 20 | 65 | Y | ML up 10 |
| 35 | 60 | Y | ML up 10 | 30 | 65 | Y | ML up 10 |
| 45 | 60 | Y | mark t/h | 40 | 65 | Y | mark t/h |
| L) | R) | 4000Hz | | | | | |
| 20 | 60 | Y | ML up 10 | | | | |
| 30 | 60 | Y | ML up 10 | | | | |
| 40 | 60 | Y | mark t/h | | | | |
| | | | | | | | |

- BC thresholds need to be obtained for 500 - 4000Hz for the left ear, which will then need to be masked (apply WHEN TO DO BC RULE, BC EAR SELECTION RULE, START LEVELS FOR BC RULE, 10dB DOWN RULE, 5dB UP RULE, 2 ASCENDING RUNS RULE, MARK THRESHOLD RULE, WHEN TO MASK BC RULE)

Apply: **INITIAL MASKING LEVEL RULE**

IML for L) ear when obtaining masked BC t/h for R) ear:

1000Hz: 5dB HL t/h + 10dB safety + 10dB occlusion effect = 25dB EM

2000Hz: 10dB HL t/h + 10dB safety = 20dB EM

4000Hz: 10dB HL t/h + 10dB safety = 20dB EM

5000Hz: 10dB HL t/h + 10dB safety + 20dB occlusion effect = 40dB EM

Apply: **PLATEAU RULE**

| Masking level (ML) dB EM | Test tone (TT) dB HL | Response | Action | Masking level (ML) dB EM | Test tone (TT) dB HL | Response | Action |
|---|---|---|---|---|---|---|---|
| L) | R) | 1000Hz | | L) | R) | 2000Hz | |
| 25 | 5 | N | TT up 5 | 20 | 10 | N | TT up 5 |
| 25 | 10 | N | TT up 5 | 20 | 15 | N | TT up 5 |
| 25 | 15 | N | TT up 5 | 20 | 20 | N | TT up 5 |
| 25 | 20 | N | TT up 5 | 20 | 25 | N | TT up 5 |
| 25 | 25 | N | TT up 5 | 20 | 30 | N | TT up 5 |
| 25 | 30 | N | TT up 5 | 20 | 35 | N | TT up 5 |
| 25 | 35 | N | TT up 5 | 20 | 40 | N | TT up 5 |
| 25 | 40 | N | TT up 5 | 20 | 45 | N | TT up 5 |
| 25 | 45 | N | TT up 5 | 20 | 50 | N | TT up 5 |
| 25 | 50 | Y | ML up 10 | 20 | 55 | Y | ML up 10 |
| 35 | 50 | Y | ML up 10 | 30 | 55 | Y | ML up 10 |
| 45 | 50 | Y | mark t/h | 40 | 55 | Y | mark t/h |
| L) | R) | 4000Hz | | L) | R) | 500Hz | |
| 20 | 10 | N | TT up 5 | 40 | 10 | N | TT up 5 |
| 20 | 15 | N | TT up 5 | 40 | 15 | N | TT up 5 |
| 20 | 20 | N | TT up 5 | 40 | 20 | N | TT up 5 |
| 20 | 25 | N | TT up 5 | 40 | 25 | N | TT up 5 |
| 20 | 30 | N | TT up 5 | 40 | 30 | N | TT up 5 |
| 20 | 35 | N | TT up 5 | 40 | 35 | N | TT up 5 |
| 20 | 40 | N | TT up 5 | 40 | 40 | Y | ML up 10 |
| 20 | 45 | N | TT up 5 | 50 | 40 | Y | ML up 10 |
| 20 | 50 | N | TT up 5 | 60 | 40 | Y | mark t/h |
| 20 | 55 | Y | ML up 10 | | | | |
| 30 | 55 | Y | ML up 10 | | | | |
| 40 | 55 | Y | mark t/h | | | | |

## Appendix III

Feedback rules developed for programming of pop-up feedback messages

**<u>Conducting Pure-tone Audiometry - RULES to guide the application of feedback</u>**

**TRANSDUCER SELECTION RULE:** Use the information obtained from otoscopy to judge what type of transducer to use. Default to insert earphones, unless there is an outer ear infection, grommets, wax or perforated ear drum, where supra aural head phones are used

**AC BEFORE BC RULE:** Air conduction thresholds are obtained first, where the stimulus or tone is presented via supra-aural headphones or insert earphones. If a hearing loss is evident (i.e. thresholds that are greater than 20dB HL) the bone conductor is used to differentiate the type of hearing loss (i.e. conductive or sensorineural)

**AC EAR SELECTION RULE**: start testing with the reported better ear; if both are the same, start with the right ear

**TEST FREQUENCY ORDER RULE**: start testing at 1000Hz, ascend up to 2000, 4000 and 8000Hz, then descend to 500 and 250Hz

**INITIAL PRESENTATION LEVEL RULE**: the level of the <u>very first tone presented</u> should be well above the patient's actual threshold so they become familiar with what they need to listen for. As a general rule, start at 30dB. If the patient doesn't respond to the first tone, increase in 15-20dB steps until they respond

**10dB DOWN RULE**: when the patient responds to a tone, decrease the signal level by 10dB, and present the tone again. Repeat descending in 10dB steps and presenting the tone until the patient doesn't respond.

**5dB UP RULE**: when the patient doesn't respond to a tone, increase by 5dB and present the tone again. Repeat ascending in 5dB steps and presenting the tone until the patient does respond

**2 ASCENDING RUNS RULE**: Mark the audiogram once the threshold has been obtained. The threshold is the level where a response is obtained for <u>two ascending runs</u> using the 10dB down, 5dB up method.

**MARK THRESHOLD RULE**: two ascending runs have been obtained. Mark the threshold at the level where the response was obtained for the two ascending runs

**COMPLETE AC EAR INFO RULE**: obtain all air conduction thresholds for the are you started with before switching to the other ear

**SUBSEQUENT PRESENTATION LEVEL RULE:** once the first threshold is determined, the start presentation level at the next frequency can be 15-20dB higher than the previous threshold

**TEST INTEROCTAVES RULE**: when noise induced hearing loss is suspected, hearing thresholds for inter-octave frequencies of 1500, 3000 and 6000Hz should be obtained. Also test inter-octaves (750, 1500, 3000, 6000Hz) if a more detailed audiogram is required, such as when a patient complains of hearing difficulties but the audiogram seems normal. The inter-octave should also be tested if the AC threshold drops 20dB between octave frequencies

**WHEN TO DO BC RULE**: Obtain bone conduction thresholds at octave frequencies from 500 to 4000Hz when air conduction thresholds at these frequencies are outside of normal limits (i.e. =/>20dB HL) to differentiate the type of hearing loss (i.e. conductive or sensorineural). Only 500 – 4000Hz can be tested reliably with bone conduction

**BC EAR SELECTION RULE**: For bone conduction, start with the worst hearing ear

**START LEVELS FOR BC RULE**: The start level for BC audiometry should be 10-15 dB above the air conduction threshold of the *better* ear

**WHEN TO MASK BC RULE**: A masked bone conduction response is required when the air conduction threshold in the test ear is greater or more severe than the unmasked bone conduction threshold (at the same frequency in the same ear), by > 10dB (i.e. a significant air-bone gap)

**MASKING TRANSDUCER RULE**: Use the information from otoscopy to judge what type of transducer to use. Default to insert earphones.

**WHAT EAR TO MASK RULE**: Narrow band masking noise is used to mask pure tones. This is played into the non-test ear via the air conduction pathway whilst true thresholds are obtained from the test ear.

| Occlusion Effect | | | | | | |
|---|---|---|---|---|---|---|
| frequency | 250Hz | 500Hz | 1000Hz | 2000Hz | 4000Hz | 8000z |
| supra-aural headphones | 30dB | 20dB | 10dB | - | - | - |
| inserts | 10dB | 10dB | - | - | - | - |

**INITIAL MASKING LEVEL RULE (BC):** The level of noise initially played into the non test ear at a given frequency is calculated with the following: level of the threshold in the non test ear + 10dB safety factor for guaranteed effective masking + occlusion effect at low frequencies

**PLATEAU RULE**: Whilst masking noise is played into the non test ear, present the tone to the test ear at the level of the unmasked threshold. Increase the level of masking noise in 10dB steps, and present the tone in the test ear until the masking has been increased by 20dB (2 steps of 10 dB) and the patient continues to respond to the tone. Increase test tone in steps of 5dB if required for it to be audible to the patient above the masking noise. The patient's masked threshold is the level where they respond to the test tone when masking noise is increased by 20dB (i.e. a 20dB plateau)

**WHEN TO MASK AC RULE:** Mask the air conduction threshold for the test ear if, at a given frequency, the threshold is =/> the best threshold of the non test ear at that frequency (air conduction or bone conduction threshold) **PLUS** the inter-aural attenuation value for the transducers delivering the air conduction signal (i.e. if there is a chance the non test ear could hear the level presented to the test ear).

| Inter-aural attenuation (IAA) | | | | | | |
|---|---|---|---|---|---|---|
| frequency | 250Hz | 500Hz | 1000Hz | 2000Hz | 4000Hz | 8000z |

| | | | | | | |
|---|---|---|---|---|---|---|
| supra-aural headphones | 40dB | 40dB | 40dB | 40dB | 40dB | 40dB |
| inserts | 75dB | 75dB | 75dB | 50dB | 50dB | 50dB |
| bone conductor | 0dB | 0dB | 0dB | 0dB | 0dB | 0dB |

**INITIAL MASKING LEVEL RULE (AC)**: The level of noise initially played into the non-test ear at a given frequency is calculated with the following: level of the threshold in the non-test ear + 10dB safety factor for guaranteed effective masking

## Appendix IV

Human Ethics Committee approval letter, participant invitation, information and consent form

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

HUMAN ETHICS COMMITTEE

Secretary, Lynda Griffioen
Email: human-ethics@canterbury.ac.nz

Ref: HEC 2012/15/LR

30 April 2012

Lynda Guard
Department of Communication Disorders
UNIVERSITY OF CANTERBURY

Dear Lynda

Thank you for forwarding your Human Ethics Committee Low Risk application for your research proposal "Feedback in a virtual patient for clinical audiology training".

I am pleased to advise that this application has been reviewed and I confirm support of the Department's approval for this project.

With best wishes for your project.

Yours sincerely

Michael Grimshaw
***Chair***
***University of Canterbury Human Ethics Committee***

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Hello,

We are offering you an opportunity to participate in research being conducted by researchers from the Communication Disorders department and the Human Interface Technology Laboratory. The project is developing a virtual client simulator to be used in clinical audiology training to supplement traditional means of teaching.

You can find more information about what taking part in the research involves in the information sheet you have been given.

Thank you for your time and consideration. It is only with the generous help of people like you that we can develop our project.

Yours sincerely,

**Alexandre Heitz**

Doctoral student

HIT Lab NZ

University of Canterbury

Email alexandre.heitz@canterbury.ac.nz

**Lynda Guard**

MAud student

Communication Disorders

University of Canterbury

Email lmg45@uclive.ac.nz

ph. 021 2633 371

**Information Sheet**

**Feedback in a Virtual Patient for Clinical Audiology Training**

---

You are invited to take part in this study as part of a project aiming to develop a 'virtual client simulator' for use as training software in the Clinical Audiology course.

**The aim of this study** is to find out the impact of feedback conditions incorporated into the virtual client simulator software on your ability to learn how to do Pure-tone Audiometry.

**Who are the researchers?**
A team of researchers from the Communication Disorders department and the Human Interface Technology laboratory is conducting this study. The researchers in the team from the Communication Disorders department are: Lynda Guard, Jonny Grady, and Dr Catherine Moran. The HIT Lab researchers are Alexandre Heitz, Dr Andreas Duenser, and Dr Christoph Bartneck. This study also forms part of Alexandre Heitz's PhD, and Lynda Guard's Masters.

**How were participants selected for this study?**
Current and prospective students in the Bachelor of Speech Pathology course have been invited to take part as they have foundation knowledge of obtaining an audiogram, yet limited practice and experience in conducting puretone audiometry.

**What will the research involve?**
We are asking you to practice pure-tone audiometry with the simulator by working through 4 cases in one block session, and then take part in assessment within 24hrs of completing these cases.

**What are the benefits of the study?**
This study will allow you to receive additional training in clinical audiology. This study will also provide information that will help in developing more realistic virtual patients and to refine our simulator before further use for clinical audiology training.

**Do I receive any incentive for completing the study?**
You will receive a $20 Westfield shopping voucher in acknowledgement of your time, and you will go in a draw to win a prize valued at $250.

**Do I have to take part?**

No, your participation is entirely voluntary (your choice). If you choose not to take part this will not affect your academic progress. We hope that you will participate because we need to obtain as many responses as possible to best refine the virtual client simulator for future use.

You may withdraw at any time. However, taking part in all activities will provide the best information for the study.

**Will my taking part in this study be kept confidential?**

Yes, the following steps have been taken to ensure the confidentiality of the research. (1) Anonymity will be maintained using aliases. (2) Access to the data is limited to the researchers named above. (3) The data will be stored securely at the University of Canterbury for five years following completion of the project and then the data will be destroyed.

**What will happen to the results of this study?**

The results of this study will allow us to refine the virtual client simulator. The results will be reported as part of the project in journal publications, conference presentations, and on the internet. It will also be reported as part of Alexandre Heitz's PhD thesis, and Lynda Guard's Master thesis.

If you would like a copy of the results of this study please contact Alexandre Heitz or Lynda Guard.

**Who has approved this study?**

This study has been reviewed and approved by the Human Interface Technology laboratory (HIT Lab NZ), the Department of Communication Disorders, and the University of Canterbury Human Ethics Committee Low Risk Approval process.

**Please contact Alexandre Heitz or Lynda Guard if you have further questions.**

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

# Consent Form

- I have read and accept my rights, and am happy to take part in this project.

- I understand that the data resulting from this study will be used by Alexandre Heitz in his doctoral research, and by Lynda Guard in her master's thesis.

- I understand that the data might appear in publications related to the virtual patient simulator project.

- I understand that the data will be held securely and kept for a minimum period of 5 years following completion of the project before being destroyed.

- I understand that my name will not be used in any presentations or reports, unless I specifically request it.

- I understand that I am able to withdraw at any time from this research.

Name: ____________________________________________

Signed: ____________________________________________

Date: ____________________________________________

## Appendix V

Basic outline of the clinical protocol for conducting pure tone audiometry with adult patients

# ADULT PURE-TONE AUDIOMETRY

This is a basic outline of the clinical protocol for conducting pure-tone audiometry with adult patients. Please read through this information which will assist practicing hearing tests with the Virtual Patient Simulator.

- **Preparation**: check referral; plan assessment procedure; get correct forms; prepare test area; perform listening checks on equipment

- **Take case history**

- **Perform ear examination & otoscopy**

- **Give instructions to patient**: indicate test purpose, what the patient will experience and what they need to do

- **Transducer selection**: choose what type of transducers (supra-aural headphones, insert earphones) based on the patient's circumstances and otoscopy findings. Default to inserts, but use supra-aural headphones if there is excessive wax, outer ear infection, grommets or a perforated ear drum

- **Transducer placement:** red for right ear, blue for left ear

- **Air Conduction (AC) test method**:
  - Start with reported <u>better</u> ear, or if both are the same, start with the <u>right</u> ear
  - Use a supra-threshold tone to familiarise patient with stimulus: testing begins at 1000Hz at 30dB HL. If they don't hear this tone (no response) increase level in steps of 15-20 dB until initial response is obtained.
  - Obtain threshold of hearing using the <u>modified Hughson-Westlake ascending method</u> (down 10/up 5):
    - decrease level in 10 dB steps if response occurs (descending)
    - increase in 5 dB steps if no response (ascending)

        - mark the threshold at the lowest level where a response occurs twice on as ascending run

    - Frequencies are tested in the following order: 1000 Hz ascend to 2, 4, and 8 kHz, descend to 500, 250 Hz.

    - Test inter-octaves frequencies (750 Hz, 1.5, 3 and/or 6Hz) if AC thresholds drop 20 dB between octave frequencies; if the person complains of hearing difficulties but the audiogram appears normal; or if noise-induced hearing loss is suspected (test 1.5, 3 & 6 kHz in this case).

Hearing is considered normal if thresholds are at or below (less than) 20 dB HL. If thresholds fall outside this range, there is a hearing loss, and the type of hearing loss (conductive – middle ear pathology, or sensorineural – cochlear pathology) can be differentiated by bone conduction testing.

- **Bone Conduction (BC) test method**:
    - Replace AC transducers with the bone conductor

    - Start with the bone conductor sitting behind the worst hearing ear

    - Frequencies tested with BC: 500 Hz, 1, 2 & 4 kHz. Only these frequencies can be tested reliably. Testing inter-octaves (1.5, 3 kHz) may add diagnostic value

    - Starting level for BC audiometry: begin 10-15 dB above the threshold of the *better* ear's AC threshold at the frequency being tested

    - Seek BC thresholds in the same manner as AC thresholds (10 dB down, 5 dB up; mark threshold at lowest level where a response occurs 2 out of 3 times on an ascending run)

- **Masking thresholds**: BC or AC thresholds need to be masked if there is any chance the opposite, or non-test ear (NTE) could hear the test tone and falsely improve the test ear's (TE) ability to hear. Masking noise (narrow band noise for pure tones) is played into the NTE to keep it busy while the true hearing threshold of the TE is confirmed.

- **When to mask**: Masking is needed in 3 situations:

    - **Obtain masked AC thresholds** when the AC threshold on the TE is =/> the AC threshold of the NTE <u>*PLUS*</u> the inter-aural attenuation (IAA) value for the transducers used **– formula: mask AC when $AC_{TE}$ > $AC_{NTE}$ + IAA**

! IAA is how much the head attenuates, or blocks the sound from the transducer travelling from one ear to the other – if a tone is played loud enough in one ear, the sound will leak to the other ear once a certain intensity is reached (see below levels) depending on the frequency of the sound and the transducer being used

| Inter-aural attenuation (IAA) | | | | | | |
|---|---|---|---|---|---|---|
| frequency | 250Hz | 500Hz | 1000Hz | 2000Hz | 4000Hz | 8000z |
| supra-aural headphones | 40dB | 40dB | 40dB | 40dB | 40dB | 40dB |
| inserts | 75dB | 75dB | 75dB | 50dB | 50dB | 50dB |
| bone conductor | 0dB | 0dB | 0dB | 0dB | 0dB | 0dB |

- 
    - **Obtain masked BC thresholds** if the AC threshold is worse (greater than) the BC threshold (same ear & frequency) by <u>15 dB or more</u>, i.e. when there is a significant air-bone gap **– formula: mask BC when $AC_{TE}$ > $BC_{TE}$ by =/>15dB**

        - NB – this only needs to be done when there is a hearing loss, i.e. AC thresholds > 20 dB HL

    - **Additional masking of AC thresholds** may be required after BC testing, if the AC threshold of the TE at a given frequency is at a level that is greater than the BC threshold of the NTE *PLUS* IAA **– formula: mask AC when $AC_{TE}$ > $BC_{NTE}$ + IA**

**! Before obtaining masked thresholds, CHECK:** Correctly identified non-test ear for receiving masking noise via AC transducer, and correctly identified test ear for confirming threshold

- **Test method for masking– "Plateau method":**

    - Calculate the initial level of the masking noise presented to the non-test ear

- **For AC threshold masking:** level of $AC_{NTE}$ threshold at the test frequency *PLUS* 10dB safety factor for guaranteed effective masking

- **For BC masking:** level of $AC_{NTE}$ threshold at the test frequency *PLUS* 10dB safety factor for guaranteed effective masking *PLUS* Occlusion effect (OE) at low frequencies

! OE is the amplification of low frequency sounds when the ear is blocked, so we account for this phenomenon by playing noise slightly louder (by the below dB values at the corresponding frequencies) into the non-test ear

| Occlusion Effect | | | | | | |
|---|---|---|---|---|---|---|
| frequency | 250Hz | 500Hz | 1000Hz | 2000Hz | 4000Hz | 8000z |
| supra-aural headphones | 30dB | 20dB | 10dB | - | - | - |
| inserts | 10dB | 10dB | - | - | - | - |

- Start masking noise in the NTE, and present the test tone to the TE at the level of the unmasked threshold you are trying to confirm

- If the patient responds, increase the level of the masking noise in 10 dB steps, and present the test tone again. Repeat this process until the masking has been increased by 20 dB and the patient continues to respond to the test tone

- If they don't respond, increase the level of the test tone in steps of 5dB, then continue to raise the masking noise in 10 dB steps

- The patient's masked threshold is the level at which they continually respond to the test tone whilst masking noise is increased by 20dB (i.e. over a 20dB plateau)

Pure-tone audiometry is complete once you have obtained all necessary AC, BC and masked thresholds.

Please don't hesitate to contact Lynda Guard if you have any questions or would like to find out more about hearing testing. Email lmg45@uclive.ac.nz.

## Appendix VI

Probes activity

**Case 1: A 53 yr old woman comes in for a hearing check. She is worried things sound a bit muffled, and she needs to turn the TV up louder than usual. Otoscopy reveals a wall of dry wax in both ears.**

1. What transducers do you select to test her hearing?

A Insert earphones
B Bone conductor
C Supra aural headphones
D Sound-field speaker
E None, the wax needs to be removed before audiometric testing can be attempted

2. What ear will you test first?

A Left
B Right
C Both, as the unmasked bone conductor be could heard in either ear
D Both, as I'm testing in the sound-field
E I wouldn't test this woman's hearing until she's had the wax removed

3. What frequency do you begin testing at?

A 1000 Hz
B 250 Hz
C 10 kHz
D 2000 Hz
E 1000 kHz

4. At what level do you first present the tone?

A -10 dB
B 20 dB
C 30 dB
D 50 dB
E 90 dB

5. She doesn't respond to the initial tone presentation. What do you do?

A Discontinue testing, as she's not hearing the tone due to the wax

B Present the tone again at the same level for 1-2 seconds

C Present the tone again at the same level, holding the presentation button down until she responds

D Increase the presentation level by 15-20 dB and present the tone for 1-2 seconds

E Increase the presentation level by 15-20 dB, present the tone, holding the presentation button down until she responds

6. She responds: how do you find her hearing threshold?

A drop in 10 dB steps until she no longer responds, then return to the initial presentation level and repeat dropping in 10 dB steps until she no longer responds. Her threshold is the level at which she no longer responds on two descending runs.

B drop in 10 dB steps until she no longer responds, go up in 5 dB steps until she responds again. Repeat dropping 10 dB and increasing 5 dB. Her threshold is where she responds two out of three times on an ascending run.

C drop in 5 dB steps until she no longer responds, go up in 10 dB steps until she responds again. Repeat dropping 5 dB and increasing 10 dB. Her threshold is where she responds two out of three times on an ascending run.

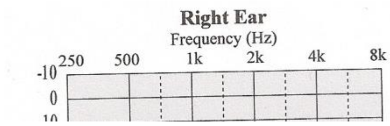
D start at -10 dB and increase in 5 dB steps until she responds. Repeat. Her threshold is where she responds on two ascending runs.

**Case 2. A 48yr old man has come in for a hearing check. He has been feeling dizzy and nauseous at times, with a sound like the roaring ocean in his right ear when he gets dizzy. He feels his hearing hasn't been that great on the right side since he's been having the dizzy spells. Otoscopy reveals clear ear canals and healthy looking eardrums on both sides.**

1. What transducers do you choose? ___________________________________________

2. What ear do you begin with? _______________________________________________

3. Write the frequencies in the order you test them:
   1) ______ 2) ______ 3) ______ 4) ______ 5) ______ 6) ______




4. At what level do you initially present the tone when testing his right ear? ______

5. At 2000Hz in the right ear, you obtain the following responses:

| Presentation level: | Response: |
|---|---|
| 60 dB HL | Yes |
| 50 dB HL | Yes |
| 40 dB HL | Yes |

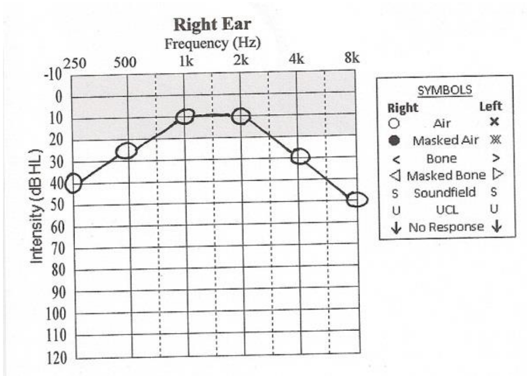
What will the next presentation level be? _________

6. At 500 Hz in the left ear, you obtain the following responses:

| Presentation level: | Response |
|---|---|
| 30 dB HL | Yes |
| 20dB HL | Yes |
| 10 dB HL | Yes |
| 0 dB HL | No |

What will the next presentation level be? _________

**Case 3. This person is having their hearing tested with insert earphones. So far you have obtained the following results:**




1. At what frequencies do you need to obtain bone conduction thresholds?

A All of them

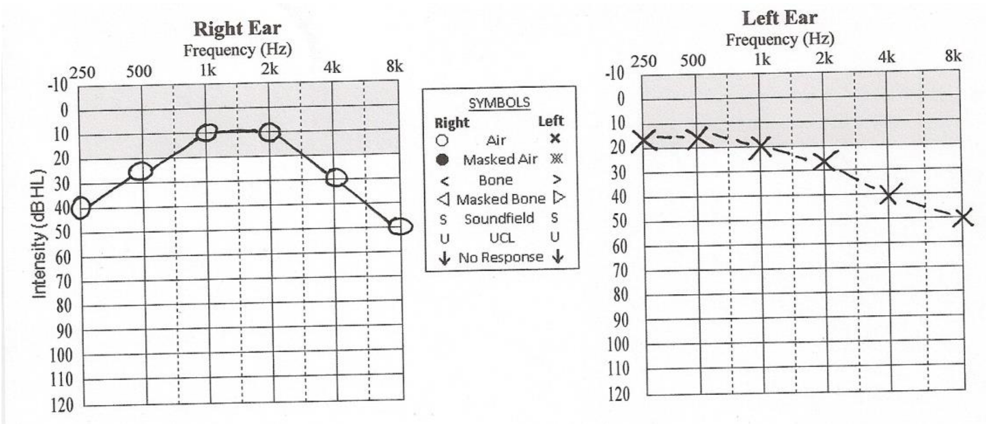
B Only those frequencies where there is a hearing loss (where AC thresholds are greater than 20 dB HL)

C At 500 Hz, 1 kHz, 2 kHz and 4 kHz

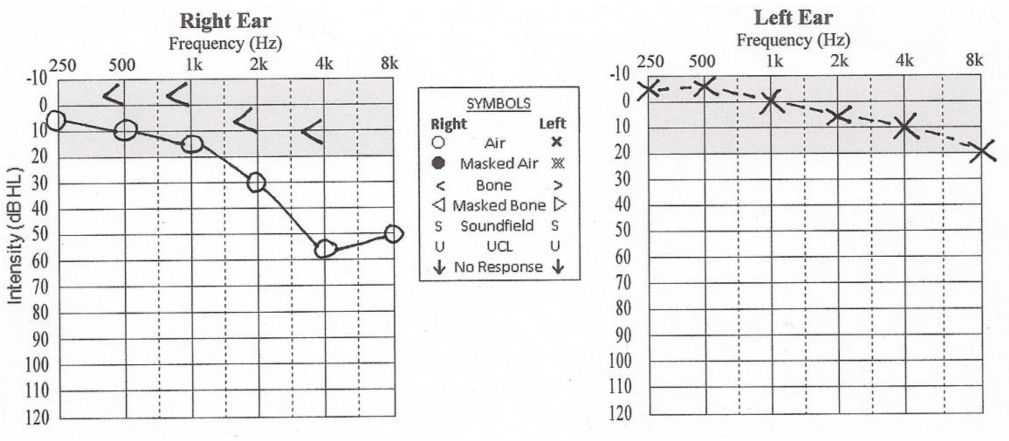
D At 500 Hz and 4 kHz

E None of them

2. Below are the air conduction results for each ear. At what frequencies do you need to obtain bone conduction thresholds in the left ear? ___________________________________



**Case 4. This person is having their hearing tested with supra-aural headphones. So far the following results have been obtained:**




1. Which bone conduction thresholds need to be masked?

A All of them
B 500 & 1000 Hz
C 1, 2 & 4 kHz
D 2 & 4 kHz
E None of them, they are in the normal range

2. You decide to mask the BC threshold at 2 kHz. What ear do you play the masking noise into?

A Both ears via the bone conductor
B Left ear only via the bone conductor
C Right ear only via the bone conductor
D Left ear only via the supra-aural headphones
E Right ear only via the supra-aural headphones

3. At what level do you start presenting the masking noise at this frequency (2kHz)?

A 5 dB HL
B 15 dB HL
C 20 dB HL
D 30 dB HL
E 40 dB HL

4. At what level do you start presenting the test tone?

A 5 dB HL
B 15 dB HL
C 20 dB HL
D 30 dB HL
E 40 dB HL

5. The patient came back for a re-test the following week. You are part way through masking their BC threshold at 2 kHz. So far you have these responses using the plateau method:

| Masking noise level: | Test signal level: | Response: | Action: |
|---|---|---|---|
| 10 dB EM | 10 dB HL | Yes | masking noise up 10 dB |
| → 20 dB EM | 10 dB HL | No | test signal up 5 dB |
| → 20 dB EM | 15 dB HL | Yes | ??? |

What is your next move?

A Mark 15 dB HL as the masked BC threshold for 2 kHz

B Leave settings as they are and present the test tone a second time to establish the threshold

C Increase the test tone by 5 dB

D Decrease the test tone by 5 dB

E Increase the masking noise level by 10 dB

6. There are some mistakes on the audiogram. What should also have been done?

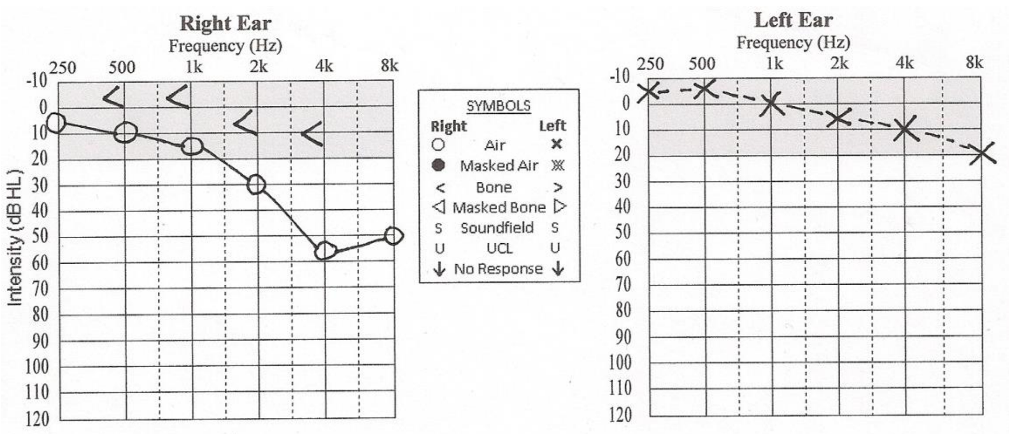
A Obtain a masked air conduction threshold for the right ear at 4 kHz and obtain unmasked bone conduction thresholds for the left ear

B Obtain an air conduction threshold at the 3 kHz inter-octave frequency on the right ear and obtain unmasked bone conduction thresholds for the left ear

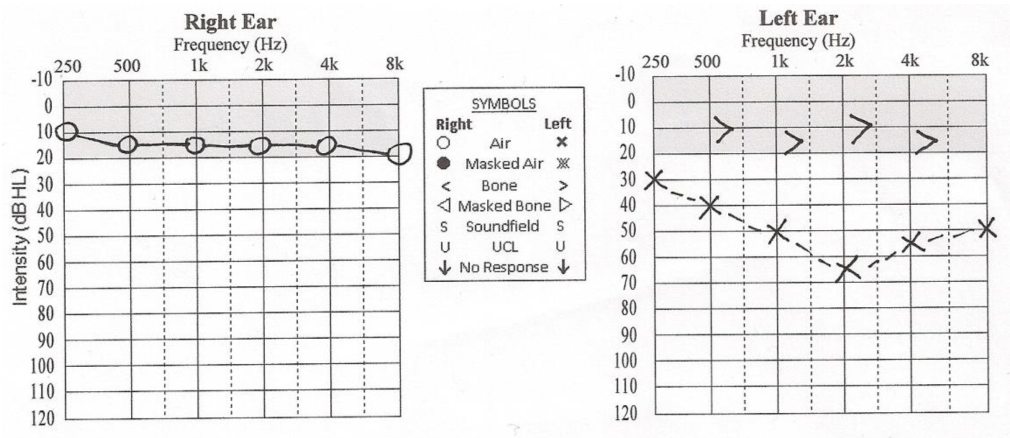
C Obtain air conduction thresholds for all the inter-octave frequencies on the right ear and obtain an unmasked bone conduction threshold at 8000 Hz for the same ear

D Obtain an air conduction threshold at the 3 kHz inter-octave frequency on the right ear and obtain a masked air conduction threshold at 4 kHz for the same ear

E Obtain unmasked bone conduction thresholds for the left ear and obtain an unmasked bone conduction threshold at 8000 Hz for the right ear



**Case 5. This person is having their hearing tested with insert earphones. So far you have obtained the following results:**




1. Which left ear bone conduction thresholds need to be masked? _____________________

2. You decide to mask the bone conduction threshold at 1 kHz in the left ear. Which ear do you play the masking noise into, with which transducer? _____________________________

3. What would the initial masking level be at this frequency? _________________

4. What intensity do you start presenting the test tone at? _____________________

5. You obtain the following responses using the plateau method when masking their BC threshold at 500 Hz.

| Masking noise level: | Test signal level: | Response: | Action: |
|---|---|---|---|
| 40 dB EM | 15 dB HL | Yes | masking noise up 10 dB |
| → 50 dB EM | 15 dB HL | Yes | masking noise up 10 dB |
| → 60 dB EM | 15 dB HL | Yes | ??? |

What is your next move? __________________________________________________

6. Is there a need to obtain any masked air conduction thresholds? If so, for which ear and which frequencies? _______________________________________

7. Is there a need to obtain any air conduction thresholds for any of the inter-octave frequencies? If so, for which ear and which frequencies? ________________________

**- - - THE END - - -**

## Appendix VII

Probes marking schedule

**Case 1: A 53 yr old woman comes in for a hearing check. She is worried things sound a bit muffled, and she needs to turn the TV up louder than usual. Otoscopy reveals a wall of dry wax in both ears.**

1. TRANSDUCER SELECTION 1 What transducers do you select to test her hearing?

A Insert earphones
B Bone conductor
C - 1 Supra aural headphones
D Sound-field speaker
E None, the wax needs to be removed before audiometric testing can be attempted

2. AC EAR SELECTION 1 What ear will you test first?

A Left
B - 1 Right
C Both, as the unmasked bone conductor be could heard in either ear
D Both, as I'm testing in the sound-field
E I wouldn't test this woman's hearing until she's had the wax removed

3. TEST FREQUENCY ORDER 1 What frequency do you begin testing at?

A - 1 1000 Hz
B 250 Hz
C 10 kHz
D 2000 Hz
E 1000 kHz

4. INITIAL PRESENTATION LEVEL 1A At what level do you first present the tone?

A -10 dB
B 20 dB
C - 1 30 dB
D - 0.5 50 dB
E 90 dB

5. INITIAL PRESENTATION LEVEL 1B She doesn't respond to the initial tone presentation. What do you do?

A Discontinue testing, as she's not hearing the tone due to the wax

B Present the tone again at the same level for 1-2 seconds

C Present the tone again at the same level, holding the presentation button down until she responds

D - 1 Increase the presentation level by 15-20 dB and present the tone for 1-2 seconds

E Increase the presentation level by 15-20 dB, present the tone, holding the presentation button down until she responds

6. T/H SEEKING METHOD 1 She responds: how do you find her hearing threshold?

A - 0.5 drop in 10 dB steps until she no longer responds, then return to the initial presentation level and repeat dropping in 10 dB steps until she no longer responds. Her threshold is the level at which she no longer responds on two descending runs.

B - 2 drop in 10 dB steps until she no longer responds, go up in 5 dB steps until she responds again. Repeat dropping 10 dB and increasing 5 dB. Her threshold is where she responds two out of three times on an ascending run.

C - 0.5   drop in 5 dB steps until she no longer responds, go up in 10 dB steps until she responds again. Repeat dropping 5 dB and increasing 10 dB. Her threshold is where she responds two out of three times on an ascending run.

D   start at -10 dB and increase in 5 dB steps until she responds. Repeat. Her threshold is where she responds on two ascending runs.

**Case 2. A 48yr old man has come in for a hearing check. He has been feeling dizzy and nauseous at times, with a sound like the roaring ocean in his right ear when he gets dizzy. He feels his hearing hasn't been that great on the right side since he's been having the dizzy spells. Otoscopy reveals clear ear canals and healthy looking eardrums on both sides.**

1. TRANSDUCER SELECTION 2 What transducers do you choose?
1 - INSERT EARPHONES, INSERTS
0.5 - EARBUDS or Supra-aurals

2. AC EAR SELECTION 2 What ear do you begin with?
1 - LEFT EAR
0.5 - BETTER EAR

3. TEST FREQUENCY ORDER 2 Write the frequencies in the order you test them:
1 – ALL CORRECT; 0.5 – 3 OR MORE CORRECT
1000 2000 4000 8000 500 250

4. INITIAL PRESENTATION LEVEL 2 At what level do you initially present the tone when testing his right ear?
2 - 50 dB HL
2 - 1: 30 dB HL; 1: "and if he doesn't hear it increase 15-20 dB"
1 - 30 dB HL

5. T/H SEEKING METHOD 2A At 2000Hz in the right ear, you obtain the following responses:

| Presentation level: | Response: |
|---|---|
| 60 dB HL | Yes |
| 50 dB HL | Yes |
| 40 dB HL | Yes |

What will the next presentation level be? 1 - 30 dB HL

6. T/H SEEKING METHOD 2B At 500 Hz in the left ear, you obtain the following responses:

| Presentation level: | Response |
|---|---|
| 30 dB HL | Yes |
| 20dB HL | Yes |
| 10 dB HL | Yes |
| 0 dB HL | No |

What will the next presentation level be? 1 - 5 dB HL

**Case 3. This person is having their hearing tested with insert earphones. So far you have obtained the following results** (pic removed)**:**

1. WHEN TO DO BC 1 At what frequencies do you need to obtain bone conduction thresholds?

A All of them
B Only those frequencies where there is a hearing loss (where AC thresholds are greater than 20 dB HL)
C At 500 Hz, 1 kHz, 2 kHz and 4 kHz
D - 1 At 500 Hz and 4 kHz
E None of them

2. WHEN TO DO BC 2 Below are the air conduction results for each ear (pic removed). At what frequencies do you need to obtain bone conduction thresholds in the left ear? 1 – 2 & 4 kHz

**Case 4. This person is having their hearing tested with supra-aural headphones. So far the following results have been obtained** (pic removed)**:**

1. WHEN TO MASK BC 1 Which bone conduction thresholds need to be masked?

A All of them
B 500 & 1000 Hz
C 1, 2 & 4 kHz
D - 1 2 & 4 kHz
E None of them, they are in the normal range

2. MASKING NOISE EAR & TRANSDUCER 1 You decide to mask the BC threshold at 2 kHz. What ear do you play the masking noise into?

A Both ears via the bone conductor
B Left ear only via the bone conductor
C Right ear only via the bone conductor
D - 1 Left ear only via the supra-aural headphones
E Right ear only via the supra-aural headphones

3. INITIAL MASKING LEVEL 1 At what level do you start presenting the masking noise at this frequency (2kHz)?

A - 0.5 5 dB HL
B - 1 15 dB HL
C 20 dB HL
D 30 dB HL
E 40 dB HL

4. INITIAL MASKED PRESENTATION LEVEL 1 At what level do you start presenting the test tone?

A - 1 5 dB HL
B 15 dB HL
C 20 dB HL
D 30 dB HL
E 40 dB HL

5. PLATEAU METHOD 1 The patient came back for a re-test the following week. You are part way through masking their BC threshold at 2 kHz. So far you have these responses using the plateau method:

| Masking noise level: | Test signal level: | Response: | Action: |
|---|---|---|---|
| 10 dB EM | 10 dB HL | Yes | masking noise up 10 dB |
| → 20 dB EM | 10 dB HL | No | test signal up 5 dB |
| → 20 dB EM | 15 dB HL | Yes | ??? |

What is your next move?

A Mark 15 dB HL as the masked BC threshold for 2 kHz

B Leave settings as they are and present the test tone a second time to establish the threshold

C Increase the test tone by 5 dB

D Decrease the test tone by 5 dB

E - 1 Increase the masking noise level by 10 dB

6. AC MASKING 1; INTEROCTAVE T/H 1 There are some mistakes on the audiogram. What should also have been done?

A - 1 Obtain a masked air conduction threshold for the right ear at 4 kHz and obtain unmasked bone conduction thresholds for the left ear

B - 1 Obtain an air conduction threshold at the 3 kHz inter-octave frequency on the right ear and obtain unmasked bone conduction thresholds for the left ear

C Obtain air conduction thresholds for all the inter-octave frequencies on the right ear and obtain an unmasked bone conduction threshold at 8000 Hz for the same ear

D - 2 Obtain an air conduction threshold at the 3 kHz inter-octave frequency on the right ear and obtain a masked air conduction threshold at 4 kHz for the same ear

E Obtain unmasked bone conduction thresholds for the left ear and obtain an unmasked bone conduction threshold at 8000 Hz for the right ear

**Case 5. This person is having their hearing tested with insert earphones. So far you have obtained the following results** (pic removed)**:**

1. WHEN TO MASK BC 2 Which left ear bone conduction thresholds need to be masked?
   1 – ALL UNMASKED BC T/Hs; 500 – 4000 Hz

2. MASKING NOISE EAR & TRANSDUCER 2 You decide to mask the bone conduction threshold at 1 kHz in the left ear. Which ear do you play the masking noise into, with which transducer?
1 – 0.5 RIGHT EAR; 0.5 INSERT

3. INITIAL MASKING LEVEL 2 What would the initial masking level be at this frequency?
   1 – 25 dB EM
   0.5 – 35 dB EM (if they are using supra-aural headphones noted in Q. 2)

4. INITIAL MASKED PRESENTATION LEVEL 2 What intensity do you start presenting the test tone at? 1 – 15 dB HL

5. PLATEAU METHOD 2 You obtain the following responses using the plateau method when masking their BC threshold at 500 Hz.

| Masking noise level: | Test signal level: | Response: | Action: |
|---|---|---|---|
| 40 dB EM | 15 dB HL | Yes | masking noise up 10 dB |
| → 50 dB EM | 15 dB HL | Yes | masking noise up 10 dB |
| → 60 dB EM | 15 dB HL | Yes | ??? |

What is your next move? 1 - Mark the masked BC threshold on the audiogram at 15 dB HL

0.5 - Stop masking; finish testing

6. AC MASKING 2 Is there a need to obtain any masked air conduction thresholds? If so, for which ear and which frequencies? 1 – 0.5 - LEFT EAR; 0.5 - 2 kHz

7. INTEROCTAVE T/H 2 Is there a need to obtain any air conduction thresholds for any of the inter-octave frequencies? If so, for which ear and which frequencies?
1 – NO

| | **SCORE** | | | | | |
|---|---|---|---|---|---|---|
| **RULE TESTED** | **MULTI CHOICE** | | **SHORT ANSWER** | | **TOTAL** | |
| TRANSDUCER SELECTION | | 1 | | 1 | | 2 |
| AC EAR SELECTION | | 1 | | 1 | | 2 |
| TEST FREQUENCY ORDER | | 1 | | 1 | | 2 |
| INITIAL PRESENTATION LEVEL | | 2 | | 2 | | 4 |
| T/H SEEKING METHOD | | 2 | | 2 | | 4 |
| WHEN TO DO BC | | 1 | | 1 | | 2 |
| WHEN TO MASK BC | | 1 | | 1 | | 2 |
| MASKING NOISE EAR & TRANSDUCER | | 1 | | 1 | | 2 |
| INITIAL MASKING LEVEL | | 1 | | 1 | | 2 |
| INITIAL MASKED PRESENTATION LEVEL | | 1 | | 1 | | 2 |
| PLATEAU METHOD | | 1 | | 1 | | 2 |
| AC MASKING | | 1 | | 1 | | 2 |
| INTEROCTAVE T/H | | 1 | | 1 | | 2 |
| **TOTAL SCORES:** | | **15** | | **15** | | **30** |

## Appendix VIII

Participant information form

## |Participant Information

ID: __________

*Please complete the following details*

**|Age:** ________ yrs

**|Gender:** M / F

**| Current study**

What degree course are you enrolled in? ☐ **BSLP** year: ________

☐ **Other** (e.g. BSc): __________

☐ **Not currently enrolled**

Have you completed CMDS 242 Introduction to Audiology? Yes / No

**|Educational Background**

Please list all previous tertiary qualifications (diplomas, degrees etc.) and date of completion:

| *Qualification* | *Year of completion* |
|---|---|
| | |
| | |
| | |

Are you considering post graduate studies in Audiology? Yes / No

**|Clinical Background**

If applicable, list all the previous clinical positions or experiences related to Audiology or hearing assessment/screening:

| *Position held/placement details* | *Duration* |
|---|---|
| | |
| | |
| | |

**|Protocol review**

How long did you spend reviewing the clinical protocol document for adult pure-tone audiometry?

☐ 0 – 15 mins ☐ +15 – 30 mins ☐ +30 – 45 mins ☐ +45 mins – 1hr ☐ +1 hr

**|Computer Skills**

How long do you use computers per week?

_______ hours

What do you use them for?

______________________________________________________________

______________________________________________________________

Do you consider that your existing skill level when using a computer is sufficient to operate the simulator? Yes / No

## Appendix IX

Usability questionnaire where participants provided feedback on their experience with the simulator

## Usability questionnaire                                    ID

*Question One - Usefulness*

a) The simulator helps me be more effective

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree                    strongly agree

b) It helps me be more productive

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree                    strongly agree

c) It is useful

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree                    strongly agree

d) It gives me more control over my learning

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree                    strongly agree

e) It makes the things I want to accomplish easier to get done

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree                    strongly agree

f) It saves me time when I use it

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree                    strongly agree

g) It meets my needs

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree                    strongly agree

h) It does everything I would expect it to do

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree                    strongly agree

*Question Two - Ease of use*

a) Overall, I am satisfied with the ease of completing the tasks in these scenarios

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree                    strongly agree

b) Overall, I am satisfied with the amount of time it took to complete the tasks in these scenarios

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree                    strongly agree

c) I can use it without written instructions

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree strongly agree

d) I don't notice any inconsistencies as I use it

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree strongly agree

e) Both occasional and regular users would like it

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree strongly agree

f) I can recover from mistakes quickly and easily

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree strongly agree

g) I can use it successfully every time

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree strongly agree

*Question Three - Ease of Learning*

a) I learned to use it quickly

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree strongly agree

b) I easily remember how to use it

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree strongly agree

c) It is easy to learn to use it

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree strongly agree

d) I quickly became skillful with it

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree strongly agree

*Question Four – Satisfaction*

a) I am satisfied with it

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree strongly agree

b) I would recommend it to a friend

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree strongly agree

c) It is fun to use

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree strongly agree

d) It works the way I want it to work

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree strongly agree

e) It is wonderful

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree strongly agree

f) I feel I need to have it

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree strongly agree

g) It is pleasant to use

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree strongly agree

*Question Five*

a) Please list five negative aspects of the simulator:

1.______________________________________________________________________

2.______________________________________________________________________

3.______________________________________________________________________

4.______________________________________________________________________

5.______________________________________________________________________

b) Please list five positive aspects of the simulator:

1.______________________________________________________________________

2.______________________________________________________________________

3.______________________________________________________________________

4.______________________________________________________________________

5.______________________________________________________________________